ELSEVIER

Bull. Sci. math. 135 (2011) 695–732

BULLETIN DES SCIENCES MATHÉMATIQUES

www.elsevier.com/locate/bulsci

Differential structure and flow equations on rough path space

Zhongmin Qian *, Jan Tudor

University of Oxford, Mathematical Institute, 24-29 St Giles', Oxford, United Kingdom

Available online 8 July 2011

Abstract

We introduce a differential structure for the space of weakly geometric p rough paths over a Banach space V for $2 < p < 3$. We begin by considering a certain natural family of smooth rough paths and differentiating in the truncated tensor series. The resulting object has a clear interpretation, even for non-smooth rough paths, which we take to be an element of the tangent space. We can associate it uniquely to an equivalence class of curves, with equivalence defined by our differential structure. Thus, for a functional on rough path space, we can define the derivative in a tangent direction analogous to defining the derivative in a Cameron–Martin direction of a functional on Wiener space. Our tangent space contains many more directions than the Cameron–Martin space and we do not require quasi-invariance of Wiener measure. In addition we also locally (globally) solve the associated flow equation for a class of vector fields satisfying a local (global) Lipshitz type condition.

MSC: 60H07

Keywords: Malliavin calculus; Rough paths; Tangent spaces

1. Introduction

The main examples of continuous random models are those constructed by solving Itô's stochastic differential equations. By means of Itô's integration, one is able to define a unique strong solution to the following Stratonovich type of differential equation

* Corresponding author.
E-mail address: qianz@maths.ox.ac.uk (Z. Qian).

0007-4497/$ – see front matter
doi:10.1016/j.bulsci.2011.07.011

$$dX_t^i = f_0^i(t, X_t)\,dt + \sum_{j=1}^{d} f_j^i(t, X_t) \circ dW_t^j, \qquad X_0 = x \tag{1.1}$$

where i runs from 1 to n, $W = (W^1, \ldots, W^d)$ is a d-dimensional Brownian motion on a probability space, and $\circ d$ denotes the Stratonovich differential. Eq. (1.1) has to be interpreted as an integration equation

$$X_t^i = x^i + \int_0^t f_0^i(s, X_s)\,ds + \sum_{j=1}^{d} \int_0^t f_j^i(s, X_s) \circ dW_s^j, \quad X_0 = x \tag{1.2}$$

where the integration is understood as the Stratonovich integrals which in turn can be converted to Itô's integrals. Suppose that the coefficients f_j^i are smooth with bounded derivatives. The important nature is that the strong solution of (1.1) is defined only almost surely, although the distribution of $X = (X_t)_{t\geqslant 0}$ is determined uniquely and is independent of the Brownian motion W. On the other hand, there is a measurable mapping F from $R^+ \times R^n \times \mathcal{C}(R^+; R^d)$ to R^n associated with (1.1) such that $X_t = F(t, x, W)$ is the unique strong solution to (1.1). Moreover, for each $t \geqslant 0$ and $\omega \in \mathcal{C}(R^+; R^d)$, $x \to F(t, x, \omega)$ is a diffeomorphism of R^n. In particular, the strong solution to (1.1) is differentiable in the initial data x, which will not be surprising to anyone who has experience with dynamical systems. It was Malliavin who first observed that the mapping $\omega \to F(t, x, \omega)$ is differential in direction h which belongs to the Cameron–Martin space of the Wiener measure, i.e. for $h \in H_0^1(R^+; R^d)$, where $H_0^1(R^+, R^d)$ is the space of all paths h in R^d whose generalized derivative $\dot{h} \in L^2(R^+; R^d)$.

In this article with the help of Lyons' continuity theorem we identify the differential structure on the space of rough paths which allows us to differentiate Wiener functionals along more tangent directions than those determined by the Cameron–Martin space. One is thus able to use the machinery of rough paths together with nonlinear functional analysis to study Wiener functionals, providing powerful mathematical tools.

In Malliavin's calculus, the Wiener functionals we are interested in are functions on the space of continuous paths $C([0, \infty); V)$, where $V = R^d$ for simplicity. The distribution μ of the standard Brownian motion in R^d is a probability measure on $C([0, \infty); V)$. If $h \in C([0, \infty); V)$ then the measurable transformation τ_h which sends a path x to $x + h$ gives rise to a push-forward measure μ_h defined by $\mu_h(A) = \mu \circ \tau_h(A)$. A classical result in probability theory says that μ_h is absolutely continuous with respect to μ if and only if h belongs to the Cameron–Martin space H consisting of all paths $h \in C([0, \infty); V)$ whose generalized derivative $\dot{h} \in L^2([0, \infty))$. Moreover, according to Cameron–Martin [1], in this case,

$$\frac{d\mu_h}{d\mu} = \exp\left[\int_0^\infty \dot{h}(t)\,d\omega(t) - \int_0^\infty \left|\dot{h}(t)\right|^2 dt \right]$$

where $d\omega(t)$ is understood as Itô's differential. This property of the Wiener measure is known as the quasi-invariance of Wiener measure.

Malliavin (see [18]) initiated a study of differentiating Wiener functionals on $C([0, \infty); V)$ in order to address the regularities of their laws. An important result is that many Wiener functionals are smooth in the Cameron–Martin directions. Because the Wiener functionals (namely solutions to some stochastic differential equations) we are interested in are only defined almost surely, it is possible to differentiate such functions on $C([0, \infty); V)$ only along the directions given in the

Cameron–Martin space, and therefore one has to perturb a path in a Cameron–Martin direction in order to preserve the measure. In rough path analysis the Wiener functionals are lifted to continuous functions on rough path space and therefore quasi-invariance of the Wiener measure is not required. This allows us to develop a calculus of variations without referring to the Wiener measure. For some interesting works in the direction of studying the existence of densities of the laws of Wiener functionals we refer the reader to the works [2,3] by Cass et al. Here, the authors use both rough path analysis as well as techniques from Malliavin calculus to show existence of densities for solutions to a class of RDE (rough differential equation).

The theory of rough paths, see [13] for a detailed discussion, was motivated in part by a desire to have a deterministic or pathwise way of dealing with stochastic differential equations. The core idea is that for paths which have infinite variation as typical stochastic paths do, for example Brownian motion, defining the integral as a Riemann sum is not sufficient. It turns out that for less regular paths, in addition to increments, one needs information about the area enclosed by a path and possibly higher order volumes in order to define an integration theory. The regularity of a rough path, in general, determines how many higher order terms must be considered. For simplicity, we restrict ourselves to the simplest true rough paths, i.e. rough paths with roughness p where $2 < p < 3$ (see below for an explanation). A rough path X with roughness p (so-called a p-rough path) is a map on the simplex $\Delta_T := \{(s, t,):\ s, t \in [0, T]\}$ taking values in the truncated tensor algebra

$$T^2(V) := 1 \oplus V \oplus V^{\otimes 2},$$

which satisfies Chen's identity, $X_{s,t} \otimes X_{t,u} = X_{s,u}$ for all $s, t, u \in [0, T]$ with $s < t < u$, and a regularity condition (1.3). Here the tensor multiplication $\otimes$ takes place in $T^2(V)$ so that

$$\begin{aligned} X^1_{s,u} &= X^1_{s,t} + X^1_{t,u}, \\ X^2_{s,u} &= X^2_{s,t} + X^2_{t,u} + X^1_{s,t} \otimes X^1_{t,u}, \end{aligned}$$

where $X^1_{s,t} \in V$, $X^2_{s,t} \in V^{\otimes 2}$ are the components of $X_{s,t}$ in V and $V^{\otimes 2}$. X has finite p-variation in the sense that

$$\sup_{\mathcal{D}} \left(\sum_l \left| X^i_{s,t} \right|^{\frac{p}{i}} \right)^{\frac{i}{p}} < \infty \tag{1.3}$$

for $i = 1, 2$.

Let $x(t) = X^1_{0t}$ for $t \leqslant T$. Then $X^1_{s,t} = x(t) - x(s)$. We sometimes say X is a rough path over the continuous path x. On the other hand, if given a continuous path x with finite variation (up to time T), one may construct a rough path X, called the canonical lift of x, by

$$\begin{aligned} X^1_{s,t} &= x(t) - x(s), \\ X^2_{s,t} &= \int_{s<u_1<u_2<t} dx(u_1) \otimes dx(u_2) \end{aligned}$$

where the integral is defined via Riemann sums. In this case Chen's identity is just the additivity of iterated integrals over different intervals. Such a p-rough path is called a *smooth* p-rough path.

The most interesting p-rough paths (where $2 < p < 3$) are of course those over Brownian motion sample paths. Observe that Brownian motion sample paths are, with probability one, Hölder continuous with exponent less than one half, which implies they have finite p-variation only for $p > 2$. It is well established that almost all Brownian motion sample paths can be lifted canonically to p-variation rough paths for $2 < p < 3$.

The space of p-rough paths equipped with the p-variation distance

$$d_p(X,Y) = \max_i \left[\sup_{\mathcal{D}} \left(\sum_l \left| X^i_{s,t} - Y^i_{s,t} \right|^{\frac{p}{i}} \right)^{\frac{i}{p}} \right]$$

is a complete metric space, denoted by $\Omega_p(V)$. This space contains two special subspaces $G\Omega_p(V)$ and $WG\Omega_p(V)$ which we define subsequently. The p-rough paths which are the limit, in p-variation distance, of a sequence of smooth rough paths are called geometric p-rough paths and denoted $G\Omega_p(V)$. While weakly geometric p-rough paths, denoted $WG\Omega_p(V)$, are the elements of $\Omega_p(V)$ that can be realized as the limit in uniform topology of the canonical lifts of bounded p-variation smooth paths. See for example [11] for more details on the differences between these spaces.

In this article, we identify a useful representation of the tangent space associated to a natural differential structure on $WG\Omega_p(V)$. The reason for our definition of derivative comes from the following observation for a finite variation path x. Given another finite variation path y, one can produce the variational path $x + \varepsilon y$ for $\varepsilon \in [0,1]$ say. This then induces a variation at the level of rough paths, $X(\varepsilon)$, of the canonical lift X of x which, due to the finite variation, is given by

$$\begin{aligned}
X(\varepsilon)^1 &= \int dx + \varepsilon \int dy, \\
X(\varepsilon)^2 &= \int dx \otimes dx + \varepsilon \left(\int dx \otimes dy + \int dy \otimes dx \right) + \varepsilon^2 \int dy \otimes dy
\end{aligned}$$

where we have suppressed the limits of integration. In this case, the derivative of $X(\varepsilon)$ as a continuous function of real numbers taking values in the linear space $C(\Delta_T, T^2(V))$ is

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} X(\varepsilon) = \left(0, \int dy, \int dx \otimes dy + \int dy \otimes dx \right).$$

Note that if x and y have finite p-variation, then the cross iterated integrals $\int dx \otimes dy$ and $\int dy \otimes dx$ have finite $\frac{p}{2}$ variation. However, in addition to varying the increment, we can also vary the second level path independently, by φ. Hence, we modify $X(\varepsilon)$ to include both first and second level variations and obtain

$$\begin{aligned}
X(\varepsilon)^1 &= \int dx + \varepsilon \int dy, \\
X(\varepsilon)^2 &= \int dx \otimes dx + \varepsilon \left(\int dx \otimes dy + \int dy \otimes dx + \varphi \right) + \varepsilon^2 \int dy \otimes dy
\end{aligned}$$

so that

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} X(\varepsilon) = \left(0, \int dy, \int dx \otimes dy + \int dy \otimes dx + \varphi \right).$$

We remark that, in some sense, $X(\varepsilon)$ is the simplest variation of X and that its derivative at 0 can be associated to the pair (Z, φ) for $Z \in \Omega_p(V \oplus V)$ with $Z^1 = (\int dx, \int dy)$,

$$Z^2 = \begin{pmatrix} \int dx \otimes dx & \int dx \otimes dy \\ \int dy \otimes dx & \int dy \otimes dy \end{pmatrix}$$

and $\varphi \in \Omega_{p/2}(V \oplus V)$. From this identification, it is possible to make rigorous the meaning of the cross iterated integrals $\int dx \otimes dy$ and $\int dy \otimes dx$ as certain projections, denoted $\pi_{12}(Z)$ and $\pi_{21}(Z)$ (see below for an explanation), of an element of $Z \in \Omega_p(V \oplus V)$ even if x and y are infinite variation paths. Hence, for each pair (Z, φ) we define a variational curve $V_{(Z,\varphi)}(\varepsilon)$ at X by

$$V_{(Z,\varphi)}(\varepsilon)^1 = X^1 + \varepsilon\pi_2(Z)^1,$$
$$V_{(Z,\varphi)}(\varepsilon)^2 = X^2 + \varepsilon\big[\pi_{12}(Z) + \pi_{21}(Z) + \varphi\big] + \varepsilon^2\pi_2(Z)^2$$

where we use the notation

$$Z = \left(1, \big(\pi_1(Z)^1, \pi_2(Z)^2\big), \begin{pmatrix} \pi_1(Z)^2 & \pi_{1,2}(Z) \\ \pi_{2,1}(Z) & \pi_2(Z)^2 \end{pmatrix}\right),$$

where π_i and π_{ij} are natural projections which should be self-evident. Finally we have to identify the equivalence classes of variations which give the same derivative. Therefore, we say that (Z, φ) is equivalent to $(\tilde{Z}, \tilde{\varphi})$ if

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} V_{(Z,\varphi)}(\varepsilon) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} V_{(\tilde{Z},\tilde{\varphi})}(\varepsilon)$$

and $\pi_1(Z) = X$ and denote the equivalence class $[Z, \varphi]$. Note that we cannot uniquely assign a variational curve to an equivalence class $[Z, \varphi]$ because we have a choice of the $\pi_2(Z)^2$ term. Our first main Theorem 15 shows that the collection of all equivalence classes $[Z, \varphi]$ is the tangent space at X, in the sense that, every possible differentiable curve of rough paths starting at X (whose derivative is taken in the function space $C([-\varepsilon, \varepsilon] : C(\Delta, T^{(2)}(V)))$) is determined uniquely by some $[Z, \varphi]$.

The idea behind the first main theorem can be described as the following. If $X(\varepsilon)$ is an arbitrary curve of rough paths, then the pair $(X^1(0), [\frac{d}{d\varepsilon}|_{\varepsilon=0} X(\varepsilon)]^1)$ will in general not have a canonical lift to $\Omega_p(V \oplus V)$. Therefore, a way of obtaining a p-rough path ($\lfloor p \rfloor = 2$) from just the increment level is required and is provided, though not uniquely, by the Lyons–Victoir extension (see [17]). The conditions for the extension theorem are the reason we are restricted to the case of weakly geometric rough paths rather than general rough paths. For the proof of this result in the general case we refer the reader to [17]. For completeness we include a proof in $\mathbb{R}^d$ for a generalized version following the same argument as in [17] in the appendix. An interesting point which distinguishes our setting from that of Malliavin calculus is that for each perturbation of a path in a Cameron–Martin direction there are infinitely many perturbations of the lifted rough path, each corresponding to a choice of "cross-iterated integrals" (projections) of the path and the Cameron–Martin direction. In other words, for each Cameron–Martin direction, there are infinitely many variations of the rough path which are not equivalent but have the same variation at the path level.

We demonstrate that the tangent space is a well-defined linear vector space, and forms a bundle over the space of rough paths, but unfortunately we do not believe that it is a fibre bundle. What is missing is the structure of local trivialization. Yet, we are still able to solve the flow equation

$$\dot{C}(\tau) = F\big(C(\tau)\big), \qquad C(0) = X$$

for $\tau \in [0, T]$ on $WG\Omega_p(V)$ for a class of functions F which are Lipschitz in some sense to be defined later (see Definition 21). Here we must consider the derivative on the left as the tangent vector $[Z, \varphi](\tau)$ uniquely associated to the curve $C(\tau)$ and F as assigning an element of the tangent space to each point of the curve. Furthermore, we say a curve $U(\cdot)$ is a solution to the flow equation if

$$\lim_{h\downarrow 0} h^{-1}\left[d_q\left(U(\tau+h), V_{[F_Z(U(\tau)),F_\varphi(U(\tau))]}(h)\right)\right]=0$$

for all τ and $U(0)=X$.

This definition requires some explanation. The reason for the appearance of d_q is that although for p-rough path space the natural metric used in the solution definition should be d_p, due to technical limitations we must use the metric d_q for some $q>p$. Specifically this is caused by the lack of an intrinsic compactness theorem for sets in Ω_p which forces us to find compactness in Ω_q. Now, loosely speaking, since V is a variational curve with parameter h for the tangent vector assigned to U at time τ, in a Banach space setting the above would reduce to

$$\lim_{h\downarrow 0} h^{-1}\left\|U(\tau+h)-\left[U(\tau)+hF\left(U(\tau)\right)\right]\right\|=0,$$

or equivalently,

$$\lim_{h\downarrow 0}\left\|\frac{U(\tau+h)-U(\tau)}{h}-F\left(U(\tau)\right)\right\|=0.$$

Therefore our definition makes sense on the rough path space and seems natural. However, the non-uniqueness of a variational curve V associated to a tangent vector means the definition may depend on the choice of variational curve. We avoid this issue by only considering a class of F such that there exists a canonical choice of variational curve associated to $F(U(\tau))$ and construct a solution via an Euler scheme.

Let us now more precisely state the main results we have briefly discussed above. The following are Theorem 15 from Section 3 and Theorem 29 from Section 4 reproduced here for convenience.

Theorem. *Let $C(\varepsilon):[-\tau,\tau]\to WG\Omega_p(V)$ be such that $C(0)=X$ and let*

$$\left(0, C'(0)^1, C'(0)^2\right)$$

be its derivative at 0. If $C'(0)^1$ has finite p-variation and $C'(0)^2$ has finite $\frac{p}{2}$-variation, there exists a unique tangent vector at X $[Z,\varphi]$ such that

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} C(\varepsilon)=\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} V_{[Z,\varphi]}.$$

That is, we identify the space of equivalence classes of curves which abstractly define the tangent space. For the precise meaning of derivative and tangent vector used in the above theorem, see Definitions 8 and 16. Furthermore, we can solve flow equations.

Theorem. *If F is a locally Lipschitz near X_0 vector field on $WG\Omega_p$, then there exists a unique solution $U:[0,\alpha]\to\Omega_q(V)$ to the flow equation for $q>p$.*

Additionally, if we strengthen the condition on F to a globally Lipschitz one then we obtain a global solution. This is the content of Theorem 31. We refer the reader to Definition 21 for a precise explanation of what we mean by "F is locally Lipschitz near X_0".

There are a great number of papers dealing with tangent vectors of Wiener space. Malliavin first introduced his idea of differentiating a functional on Wiener space in [18]. Since then, many topics have been developed and many powerful techniques have been produced within his calculus. In particular, much work has been done extending the Cameron–Martin quasi-invariance theorem to Wiener space on a based manifold as a corollary to the existence of a flow associated to some vector fields. Many of these papers consider more general tangent vectors of Wiener space.

In [5], Cruzeiro considered vector fields as maps from Wiener space to the Cameron–Martin space, which satisfied certain exponential integrability estimates. For these special vector fields, the author was able to produce a solution flow by approximating the fields through projection onto a finite-dimensional space which was identified with $\mathbb{R}^n$ for which ordinary differential equation techniques produce a flow. As a corollary, it was obtained that for any constant vector field (i.e. for all $x \in C([0,1]; R)$, $F(x) = h$ for some h in Cameron–Martin space) the measure induced by the corresponding flow $U_t(x)$ is absolutely continuous with respect to the Wiener measure. This paper instigated a series of works related to extending the results to manifolds. In [19,20], the authors showed that the Wiener measure on loops over a compact connected Lie group is quasi-invariant with respect to the left action of paths on the group which have finite energy, i.e. $\int_0^1 \|u^{-1}(t)\dot{u}(t)\|^2\,dt < \infty$. However, for the right action in [19], Malliavin deals with the Wiener measure on a connected Lie group of matrices, for which a negative result is obtained. He defines the tangent space as the space of continuous paths u taking values in the Lie algebra such that $\int_0^1 \|\dot{u}(t)\|^2\,dt < \infty$. The main theorem then shows that if the adjoint representation is not unitary, one cannot obtain quasi-invariance of the measure induced by infinitesimal right action by elements of this tangent space.

A major step in extension to manifolds was accomplished by Driver in [7] where the author proves a quasi-invariance result for the path space of a compact manifold without boundary, thereby extending the work of Cruzeiro. He defines a tangent vector field on the based path space W as a map $X^h(\omega)(s) := H(\omega)(s)h(s)$ where $H(\omega)(s)$ is the stochastic parallel translation along ω on the interval $[0,s]$, and $h : [0,1] \to T_oM$ with $h(0) = o$ and h has finite energy. The flow is constructed through geometric means whenever the covarient derivative satisfies the torsion skew symmetric condition. Furthermore, the quasi-invariance of the induced measure is proved. Note that in the case that $M = \mathbb{R}^n$, the tangent vector fields reduce to $X^h(\omega)(s) = h(s)$, i.e. the usual Cameron–Martin space if we identify the tangent space with $\mathbb{R}^n$. In this case, the flow is solved to be $u(t) = \omega + th$. The torsion skew symmetric condition was relaxed in [14] to allow for any affine connection which is adjoint skew symmetric (if the affine connection preserves the Riemannian metric then the adjoint skew symmetric and torsion skew symmetric conditions are equivalent). An extension of these results is made in [8] where a flow is produced and a quasi-invariance theorem is shown for the case of vector fields defined as above, but where h is replaced by a continuous semi-martingale of particular form. In addition, Elworthy and Li [9,10] have developed an approach using solution maps of stochastic differential equations to construct the (Bismut) tangent space.

A deterministic construction of Driver's flow on a closed Riemannian manifold is produced in [15] using the theory of rough paths. Here Lyons and Qian construct a flow for a class of vector fields obtained from solving a class of rough differential equations. They apply this to the construction of Driver's flow using the fact that one can solve the flow equation for a geometric vector field by considering the solution flow of an Ito map obtained by solving a certain differential equation.

A further notion of tangent spaces was investigated by Cipriano, Cruziro, and Malliavin in [4,6]. In these works, the authors develop the notion of a process tangent to Wiener space defined to be an $\mathbb{R}^d$ valued semi-martingale ξ satisfying the Itô equation $d\xi_i(t) = A_{ij}\,dx_j(t) + B_j\,dt$ where the anti-symmetric matrix coefficients A_{ij} are semi-martingales which also have a Stratonovich representation. An associated flow is also constructed and shown to have a push forward measure which is absolutely continuous with respect to the Wiener measure with density in $L^p(d\mu)$ for all $p \geqslant 1$. Of particular interest is the fact that a tangent process has a representation as the solution to the equation $\eta_t = \int_0^t \eta_t\,dx(t) + \gamma\,dt$. In some sense, this is related to the information contained in cross iterated integrals of a process with the Brownian path together with some additive function. This information is contained in our definition of the tangent in a deterministic way.

Yet another approach to defining derivatives of functionals on the Wiener space is considered in [12]. Instead of considering the variation $F(x+th)$, the authors consider a class of measure preserving transformations T_t giving associated variation $F(T_t x + th)$ and construct a solution to the related flow equation.

We stress that in the above non-rough path approaches, the vector fields are limited to a subspace of the Wiener space (usually the Cameron–Martin space) and the flow is defined for almost every (with respect to Wiener measure) element of continuous path space. In our definition we provide a much larger class of tangent directions which is defined point wise without reference to a measure. In fact, given any element of the same rough path space, there are infinitely many tangent directions which in some sense are variations in the direction of the given rough path.

The paper is organized as follows. In Section 2 we formally present the machinery we require from rough path theory with the exception of the Lyons–Victoir extension theorem which is instead discussed in Appendix A. Section 3 is devoted to the definition and properties of the tangent space while the final Section 4 covers the construction of the local and global flows.

2. Preliminaries

Let us discuss the tools we will need from rough path theory. Historically the theory of rough paths was developed as an approach to making deterministic sense of differential equations of the type

$$dx_t = F(x_t)\,dy_t,$$
$$x_0 = \xi$$

where the path y is very irregular in time parameter t. In the case that y is not differentiable, solutions x must interpreted as an integral

$$x_{\cdot} = x_0 + \int_0^{\cdot} F(x_t)\,dy_t$$

and if y is regular enough to have bounded variation, then ODE theory tells us that, for Lipschitz F, the above equation has a bounded variation solution. Rough path theory provides an extension of the classical theory significantly beyond bounded variation (Young integration can be seen as a simpler extension). In order to go farther, some more information than the path increments is required to define the integral. And this was conjectured in some sense by Föllmer (see the history section of Lyon's original paper [13]). More specifically, he guessed that knowing the increment and Levy area would be sufficient to solve stochastic differential equations. In fact, Lyons showed that a deterministic approach using just the increments is not possible. With this in mind and the fact that higher order iterated integrals contain area information, we want for paths of finite p-variation, we want to be able to define the iterated integrals

$$\int\limits_{s<t_1<\cdots<t_n<t} dy_{t_1} \otimes \cdots \otimes dy_{t_n}.$$

Notice that if y has finite variation, then the first iterated integral is just the increment $y_t - y_s$ and the second iterated integral is $\lim_{m(\mathcal{D})\to 0} \sum_l (y_{t_l} - y_s) \otimes (y_t - y_{t_l})$ so in this case, the higher order term is determined by the increment. Also, since up to this point we don't have an integration theory for rough paths, such objects are not defined for p-variation paths for $p \geqslant 2$. Instead, we consider a p-rough path as an object in $\bigoplus_{i=i}^{\lfloor p \rfloor} V^{\otimes i}$ where the element in $V^{\otimes i}$ behaves algebraically like an ith order iterated integral and satisfies a finite p-variation condition. The reason we consider elements only up to order $\lfloor p \rfloor$ is that for any p-rough path, there exists a unique extension to $\bigoplus_{i=i}^{\infty} V^{\otimes i}$ which has finite p-variation (see for example the extension Theorem 3.1.2 in [16]).

Let us now give the formal framework for considering path increments as well as higher order elements.

Definition 1. Let V be a Banach space and $T^n(V) = \bigoplus_{i=0}^{n} V^{\otimes i}$ be its tensor powers. A continuous function $X : \Delta_T \to T^n(V)$ denoted at each pair (s,t) by

$$X_{s,t} = \left(X^0_{s,t}, \ldots, X^n_{s,t}\right)$$

is said to be a multiplicative functional of degree n in V if

$$X_{s,t} \otimes X_{t,u} = X_{s,u} \tag{2.1}$$

for all $s,t,u \in [0,T]$ satisfying $s<t<u$, where the tensor product is taken in $T^n(V)$.

The algebraic condition (2.1), referred to throughout as Chen's identity, captures the additivity property of integrals over regions. Indeed, for x a path with bounded variation, letting $X^2_{s,u} = \int_{s<t_1<t_2<u} dx_{t_1} \otimes dx_{t_2}$, we have

$$\begin{aligned}
\int\limits_{s<t_1<t_2<u} dx_{t_1} \otimes dx_{t_2} &= \int_s^t (x_{t_2} - x_s) \otimes dx_{t_2} + \int_t^u (x_{t_2} - x_t) \otimes dx_{t_2} + \int_t^u (x_t - x_s) \otimes dx_{t_2} \\
&= \int\limits_{s<t_1<t_2<t} dx_{t_1} \otimes dx_{t_2} + \int\limits_{t<t_1<t_2<u} dx_{t_1} \otimes dx_{t_2} \\
&\quad + (x_t - x_s) \otimes (x_u - x_t) \\
&= X^2_{s,t} + X^2_{t,u} + X^1_{s,t} \otimes X^1_{t,u} \\
&= X_{s,t} \otimes X_{t,u}.
\end{aligned}$$

Finally, we form a p-rough path by the imposition of the analytic finite p-variation condition on X.

Definition 2. A p-rough path in V is a multiplicative functional of degree $\lfloor p \rfloor$ in V with finite p-variation, i.e.

$$\sup_{\mathcal{D}\subseteq[0,T]} \sum_{\mathcal{D}} \| X^i_{t_l,t_{l+1}} \|_V^{\frac{p}{i}} < \infty$$

for each $i \in \{1, \dots, \lfloor p \rfloor\}$. The space of all p-rough paths in V is denoted by $\Omega_p(V)$ and can be equipped with the distance

$$d_p(X, Y) := \max_{i\in\{1,\dots,\lfloor p\rfloor\}} \sup_{\mathcal{D}\subseteq[0,T]} \left(\sum_{\mathcal{D}} \| X^i_{t_l,t_{l+1}} \|_V^{\frac{p}{i}} \right)^{\frac{i}{p}}$$

in which case it is a complete metric space.

There are two special spaces of rough paths. The p-rough paths which are the limit, in p-variation distance, of a sequence of smooth rough paths are called geometric p-rough paths and denoted $G\Omega_p(V)$. And the elements of $\Omega_p(V)$ that can be realized as the limit in uniform topology of the canonical lifts of bounded p-variation smooth paths are called weakly geometric p-rough paths and are denoted $WG\Omega_p(V)$. The strict inclusions

$$G\Omega_p(V) \subset WG\Omega_p(V) \subset \Omega_p(V)$$

hold.

The following is standard and shows that up to reparametrization, p-rough paths are closely related to $\frac{1}{p}$-Hölder continuous paths.

Proposition 3. *Let $X \in \Omega_p(V)$ and assume X^1 is not zero on any interval. Also define $\tau : [0, T] \to \mathbb{R}^+$ by*

$$\tau(t) = \frac{\omega(t)T}{\omega(T)}$$

where $\omega(t)$ is the pth power of the p-variation of X up to time t on the path level, that is

$$\omega(t) = \sum_{i=1}^{2} \sup_{\mathcal{D}\subseteq[0,t]} \sum_{\mathcal{D}} \left| X^i_{t_{l-1}t_l} \right|^{\frac{p}{i}}.$$

Then

$$\left| X^i_{\tau^{-1}(s),\tau^{-1}(t)} \right| \leqslant \left(\frac{\omega(T)}{T} \right)^{\frac{i}{p}} (t-s)^{\frac{i}{p}}.$$

Proof. We have by the sub-additivity of p-variation over subintervals,

$$\begin{aligned}
\left| X^i_{\tau^{-1}(s)\tau^{-1}(t)} \right|^{\frac{p}{i}} &\leqslant \sum_{i=1}^{2} \sup_{\mathcal{D}\subseteq[\tau^{-1}(s),\tau^{-1}(t)]} \sum_{\mathcal{D}} \left| X^i_{t_{l-1}t_l} \right|^{\frac{p}{i}} \\
&\leqslant \sum_{i=1}^{2} \sup_{\mathcal{D}\subseteq[0,\tau^{-1}(t)]} \sum_{\mathcal{D}} \left| X^i_{t_{l-1}t_l} \right|^{\frac{p}{i}} - \sum_{i=1}^{2} \sup_{\mathcal{D}\subseteq[0,\tau^{-1}(s)]} \sum_{\mathcal{D}} \left| X^i_{t_{l-1}t_l} \right|^{\frac{p}{i}} \\
&= \omega\big(\tau^{-1}(t)\big) - \omega\big(\tau^{-1}(s)\big).
\end{aligned}$$

Then by the definition of τ,

$$\omega(t)=\frac{\omega(T)}{T}\tau(t)$$

so plugging in τ^{-1},

$$\omega\big(\tau^{-1}(t)\big)=\frac{\omega(T)}{T}t$$

and hence

$$\big|X^i_{\tau^{-1}(s)\tau^{-1}(t)}\big|^{\frac{p}{i}}\leqslant\frac{\omega(T)}{T}(t-s).$$

Finally, taking the pth root gives the result. □

The next bound will be used to prove the well-known extrinsic compactness result for sets in $\Omega_p(V)$.

Lemma 4. *Let $X,Y\in\Omega_p(V)$. Then for $q>p$ we have the following bound*

$$d_q(X,Y)\leqslant Cd_p(X,Y)^{\frac{p}{q}} \tag{2.2}$$

where

$$C=\max\Big\{\Big(2\sup_t\big|X^1_{0,t}-Y^1_{0,t}\big|\Big)^{\frac{q-p}{q}},\\ \Big(2\sup_t\big|X^2_{0,t}-Y^2_{0,t}\big|\Big(1+2\Big(\sup_t\big|X^1_{0,t}\big|+\sup_t\big|Y^1_{0,t}\big|\Big)\Big)\Big)^{\frac{q-p}{q}}\Big\}.$$

Proof. First we make the elementary estimate

$$\begin{aligned}&\sup_{\mathcal{D}}\Big(\sum_l\big|X^i_{t_lt_{l+1}}-Y^i_{t_lt_{l+1}}\big|^{\frac{q}{i}}\Big)^{\frac{i}{q}}\\&\quad=\sup_{\mathcal{D}}\Big(\sum_l\big|X^i_{t_lt_{l+1}}-Y^i_{t_lt_{l+1}}\big|^{\frac{q-p}{i}}\big|X^i_{t_lt_{l+1}}-Y^i_{t_lt_{l+1}}\big|^{\frac{p}{i}}\Big)^{\frac{i}{q}}\\&\quad\leqslant\Big(\sup_{s,t\in\Delta_T}\big|X^i_{s,t}-Y^i_{s,t}\big|\Big)^{\frac{q-p}{q}}\sup_{\mathcal{D}}\Big(\sum_l\big|X^i_{t_lt_{l+1}}-Y^i_{t_lt_{l+1}}\big|^{\frac{p}{i}}\Big)^{\frac{i}{q}}.\end{aligned}$$

Then, by Chen's identity,

$$\begin{aligned}\big|X^1_{s,t}-Y^1_{s,t}\big|&=\big|X^1_{0,t}-X^1_{0,s}-\big(Y^1_{0,t}-Y^1_{0s}\big)\big|\\&\leqslant2\sup_t\big|X^1_{0,t}-Y^1_{0,t}\big|\end{aligned}$$

and

$$\begin{aligned}\big|X^2_{s,t}-Y^2_{s,t}\big|&=\big|X^2_{0,t}-X^2_{0,s}-X^1_{0,s}\otimes X^1_{s,t}-\big(Y^2_{0,t}-Y^2_{0,s}-Y^1_{0,s}\otimes Y^1_{s,t}\big)\big|\\&\leqslant2\sup_t\big|X^2_{0,t}-Y^2_{0,t}\big|+\big|\big(X^1_{0,s}-Y^1_{0,s}\big)\otimes X^1_{s,t}+Y^1_{0,s}\otimes\big(X^1_{s,t}-Y^1_{s,t}\big)\big|\\&\leqslant2\sup_t\big|X^2_{0,t}-Y^2_{0,t}\big|+4\sup_t\big|X^1_{0,t}-Y^1_{0,t}\big|\Big(\sup_t\big|X^1_{0,t}\big|+\sup_t\big|Y^1_{0,t}\big|\Big)\end{aligned}$$

which gives the result. □

We will use the following compactness result in the proof of the existence of flow equation solutions on the space of rough paths. The fact that it is not intrinsic is the source of the awkward fact that our solution of a flow equation on $WG\Omega_p$ lives in Ω_q.

Theorem 5. *If V is finite-dimensional then any $\mathcal{A} \subset \Omega_p(V)$ satisfying*

$$\sup_{\mathcal{A}} d_p(0, X) \leqslant M$$

is relatively compact.

Proof. Consider the family of paths in V given by $\{X^1_{0,\cdot}\colon X \in \mathcal{A}\}$. The uniformly bounded p-variation implies

$$\sup_{\mathcal{A}} \left|X^1_{s,t}\right| \leqslant M$$

so that because V is finite-dimensional, for fixed t $\{X^1_{0,t}\colon X \in \mathcal{A}\}$ is relatively compact. Also, by Proposition 3

$$\begin{aligned}\left|X^1_{0,t} - X^1_{0,s}\right| &= \left|X^1_{s,t}\right| \\ &\leqslant C|t-s|^{\frac{1}{p}}\end{aligned}$$

where the constant depends only on the length of the time interval and the total p-variation of X. So we have shown equicontinuity. Hence, the Ascoli–Arzela theorem implies that $\{X^1_{0,\cdot}\colon X \in \mathcal{A}\}$ is relatively compact in the uniform topology on $C([0,T], V)$.

Now, let $X(n)$ be any sequence in $\mathcal{A}$. Then by the above argument there exists a subsequence $X(n_k)$ such that $X^1_{0,\cdot}(n_k)$ converges in sup norm. Next consider the family of paths in $V^{\otimes 2}$ given by $\{X^2_{0,\cdot}(n_k)\}$. Again we have pointwise relative compactness in $V^{\otimes 2}$ due to finite dimensionality and uniformly bounded p-variation. Also, just as above, we apply Chen's identity and Proposition 3 to arrive at

$$\begin{aligned}\left|X^2_{0,t}(n_k) - X^2_{0,s}(n_k)\right| &= \left|X^2_{s,t}(n_k) + X^1_{0,s}(n_k) \otimes X^1_{s,t}(n_k)\right| \\ &\leqslant \left|X^2_{s,t}(n_k)\right| + \left|X^1_{0,s}(n_k)\right|\left|X^1_{s,t}(n_k)\right| \\ &\leqslant C\left[(t-s)^{\frac{2}{p}} + s^{\frac{1}{p}}(t-s)^{\frac{1}{p}}\right]\end{aligned}$$

for C independent of k. Then we have equicontinuity of the paths $X^2_{0,\cdot}$ as

$$|t-s| \leqslant \min\left\{\left(\frac{\varepsilon}{2C}\right)^{\frac{p}{2}}, \left(\frac{\varepsilon}{2CT^{\frac{1}{p}}}\right)^{p}\right\}$$

implies $|X^2_{0,t}(n_k) - X^2_{0,s}(n_k)| \leqslant \varepsilon$. Therefore, after a second application of the Ascoli–Arzela theorem we can extract a further subsequence such that both $X^1_{0,\cdot}(n_{k_l})$ and $X^2_{0,\cdot}(n_{k_l})$ converge in the uniform norm on V and $V^{\otimes 2}$ respectively.

Finally, consider any sequence in $\mathcal{A}$. The above allows us to extract a subsequence such that each $X^i_{0,\cdot}(l)$ converges in uniform norm $V^{\otimes i}$. Then the bound (2.2) together with the uniformly bounded p-variation implies the subsequence is also Cauchy in $\Omega_q(V)$. Therefore completeness of $\Omega_q(V)$ gives the result. $\square$

Let us now introduce some notation and basic operations on rough paths. We will often consider rough paths which at the first tensor level consist of a pair of rough paths, i.e. rough paths in $V \oplus W$, so we use the following notation for the component parts of such rough paths Z. We use

$$Z^1 = \big(\pi_1(Z)^1, \pi_2(Z)^1\big),$$
$$Z^2 = \begin{pmatrix} \pi_1(Z)^2 & \pi_{1,2}(Z) \\ \pi_{2,1}(Z) & \pi_2(Z)^2 \end{pmatrix}$$

due to the decomposition

$$T^2(V \oplus W) = 1 \oplus (V \oplus W) \oplus \big(V^{\otimes 2} \oplus (V \otimes W) \oplus (W \otimes V) \oplus W^{\otimes 2}\big),$$

and may also denote Z by $(\pi_1(Z), \pi_2(Z))$ where the projections

$$\pi_1(Z) = \big(1, \pi_1(Z)^1, \pi_1(Z)^2\big),$$
$$\pi_2(Z) = \big(1, \pi_2(Z)^1, \pi_2(Z)^2\big)$$

are p-rough paths in V and W respectively. In this case, by identifying $\oplus$ with $+$ when the objects are in the same space, Chen's identity (2.1) is equivalent to

$$\pi_i(Z)^1_{s,u} = \pi_i(Z)^1_{s,t} + \pi_i(Z)^1_{t,u},$$
$$\pi_i(Z)^2_{s,u} = \pi_i(Z)^2_{s,t} + \pi_i(Z)^2_{t,u} + \pi_i(Z)^1_{s,t} \otimes \pi_i(Z)^1_{t,u},$$
$$\pi_{ij}(Z)_{s,u} = \pi_{ij}(Z)_{s,t} + \pi_{ij}(Z)_{t,u} + \pi_i(Z)^1_{s,t} \otimes \pi_j(Z)^1_{t,u}.$$

If we know two rough paths X and Y as a single rough path Z in $V \oplus V$, i.e. in some sense we know their cross iterated integrals, then we may add the projections of Z in the following sense.

Proposition 6. *Let $Z = (X, Y)$ be a p-rough path in $V \oplus V$. Then*

$$\big(1, X^1 + Y^1, X^2 + \pi_{1,2}(Z) + \pi_{2,1}(Z) + Y^2\big)$$

is a p-rough path in V.

Proof. This is from Chen's identity for rough paths in $V \oplus V$ given above in Eq. (2.1). □

Scalar multiplication is well defined for rough paths in the following sense.

Proposition 7. *Let X be a p-rough path in V and for all $\lambda \in \mathbb{R}$ define λX by*

$$(\lambda X)^1 = \lambda X^1,$$
$$(\lambda X)^2 = \lambda^2 X^2.$$

Then, λX is a p-rough path over V.

Proof. Immediate from Chen's identity and the properties of scalar multiplication on tensor product spaces. □

3. Tangent space construction

As the space of rough paths is a nonlinear subset of the linear vector space $C(\Delta_T; T^2(V))$, its tangent space is defined to consist of equivalence classes of curves $C : [-\tau, \tau] \to \Omega_p(V)$ which have equal derivative in the following sense.

Definition 8. The derivative of a map $C : [-\tau, \tau] \to C(\Delta_T; T^2(V))$ at ε is defined by the relation

$$C'(\varepsilon) := \lim_{h \to 0} \frac{C(\varepsilon + h) - C(\varepsilon)}{h}$$

where the limit is taken in the standard uniform topology on the vector space $C([-\tau, \tau] : C(\Delta_T : T^2(V)))$. We will also use the notation $\frac{d}{d\varepsilon} C(\varepsilon)$ to denote the derivative at ε.

We emphasize that the derivative at ε is in general no longer a rough path because Chen's identity does not hold. Our goal in this section is to construct a useful representation of the abstract notion of equivalence classes of curves with equal derivative, in terms of rough paths. In order to investigate the structure of these equivalence classes, we being by considering curves of a special form which can be thought of as analogous to real valued curves of the form $x + \epsilon y$ for fixed x and y with varying ϵ. As rough path space is not linear, it is not immediately clear how one should vary X as addition of rough paths does not make sense. Even with this fact, we can define a kind of addition of rough paths X and Y if we have some information not contained in just X and Y. That is, if we have $Z \in \Omega_p(V \oplus V)$ with $\pi_1(Z) = X, \pi_2(Z) = Y$. Naturally, there are many such Z that give rise to X and Y through the projections, but they give different "sums" as the definition of the "sum" depends on cross iterated integrals or the projections $\pi_{1,2}(Z)$ and $\pi_{2,1}(Z)$. This tells us that the information in Y alone is not sufficient to determine a direction in rough path sense. Yet Y together with the cross iterated integrals of X and Y is enough to determine the sum of X and Y in rough path sense. This suggests as a first step we consider elements Z of $\Omega_p(V \oplus V)$ such that $\pi_1(Z) = X$. For such Z we can "add" $\pi_1(Z)$ and $\pi_2(Z)$ and form a new rough path K in the same space as each of the projections according to the formula

$$K^1_{s,t} := \pi_1(Z)^1_{s,t} + \pi_2(Z)^1_{s,t}, \tag{3.1}$$

$$K^2_{s,t} := \pi_1(Z)^2_{s,t} + \pi_2(Z)^2_{s,t} + \pi_{12}(Z)_{s,t} + \pi_{21}(Z)_{s,t} \tag{3.2}$$

where

$$\begin{aligned} Z_{s,t} &= \left(1, Z^1_{s,t}, Z^2_{s,t}\right) \\ &= \left(1, \pi_1(Z)^1_{s,t}, \pi_2(Z)^1_{s,t}, \pi_1(Z)^2_{s,t}, \pi_{12}(Z)_{s,t}, \pi_{21}(Z)_{s,t}, \pi_2(Z)^2_{s,t}\right). \end{aligned}$$

Then,

$$\begin{aligned} K^1_{s,t}(\varepsilon) &:= \pi_1(Z)^1_{s,t} + \varepsilon \pi_2(Z)^1_{s,t}, \\ K^2_{s,t}(\varepsilon) &:= \pi_1(Z)^2_{s,t} + \varepsilon\left(\pi_{12}(Z)_{s,t} + \pi_{21}(Z)_{s,t}\right) + \varepsilon^2 \pi_2(Z)^2_{s,t} \end{aligned}$$

is the rough paths equivalent of adding ε times a direction to X.

Proposition 9. *If $Z \in \Omega_p(V \oplus V)$ is of the form*

$$\begin{aligned} Z^1_{s,t} &= \left(\pi_1(Z)^1, \pi_2(Z)^1\right), \\ Z^2_{s,t} &= \begin{pmatrix} \pi_1(Z)^2_{s,t} & \pi_{12}(Z)_{s,t} \\ \pi_{21}(Z)_{s,t} & \pi_2(Z)^2_{s,t} \end{pmatrix} \end{aligned}$$

then the path $V_{Z,2}(\varepsilon)$ defined by

$$V_{Z,2}(\varepsilon)^1_{s,t} = \left(\pi_1(Z)^1_{s,t}, \varepsilon\pi_2(Z)^1_{s,t}\right),$$
$$V_{Z,2}(\varepsilon)^2_{s,t} = \begin{pmatrix} \pi_1(Z)^2_{s,t} & \varepsilon\pi_{12}(Z)_{s,t} \\ \varepsilon\pi_{21}(Z)_{s,t} & \varepsilon^2\pi_2(Z)^2_{s,t} \end{pmatrix}$$

is also in $\Omega_p(V \oplus V)$. In particular, this implies the function $V_Z(\varepsilon)$ defined by

$$V_Z(\varepsilon)^1_{s,t} = \pi_1(Z)^1_{s,t} + \varepsilon\pi_2(Z)^1_{s,t},$$
$$V_Z(\varepsilon)^2_{s,t} = \pi_1(Z)^2_{s,t} + \varepsilon\pi_{12}(Z)_{s,t} + \varepsilon\pi_{21}(Z)_{s,t} + \varepsilon^2\pi_2(Z)^2_{s,t}$$

is in $\Omega_p(V)$.

Proof. We must verify that Chen's identify is satisfied so we examine $V_Z(\varepsilon)_{s,t} \otimes V_Z(\varepsilon)_{tu}$. For simplicity let X and Y denote $\pi_1(Z)$ and $\pi_2(Z)$ respectively. By definition,

$$\begin{aligned}&\left(V_{Z,2}(\varepsilon)_{s,t} \otimes V_{Z,2}(\varepsilon)_{t,u}\right)^2 \\ &\quad= \begin{pmatrix} X^2_{s,t} + X^2_{t,u} + X^1_{s,t} \otimes X^1_{t,u} & \varepsilon(\pi_{1,2}(Z)_{s,t} + \pi_{1,2}(Z)_{t,u} + X^1_{s,t} \otimes Y^1_{t,u}) \\ \varepsilon(\pi_{2,1}(Z)_{s,t} + \pi_{2,1}(Z)_{t,u} + Y^1_{s,t} \otimes X^1_{t,u}) & \varepsilon^2(Y^2_{s,t} + Y^2_{t,u} + Y^1_{s,t} \otimes Y^1_{t,u}) \end{pmatrix}.\end{aligned}$$

But since Z is multiplicative, we have

$$\begin{aligned}&(Z_{s,t} \otimes Z_{t,u})^2 \\ &\quad= \begin{pmatrix} X^2_{s,t} + X^2_{t,u} + X^1_{s,t} \otimes X^1_{t,u} & \pi_{1,2}(Z)_{s,t} + \pi_{1,2}(Z)_{t,u} + X^1_{s,t} \otimes Y^1_{t,u} \\ \pi_{2,1}(Z)_{s,t} + \pi_{2,1}(Z)_{t,u} + Y^1_{s,t} \otimes X^1_{t,u} & Y^2_{s,t} + Y^2_{t,u} + Y^1_{s,t} \otimes Y^1_{t,u} \end{pmatrix} \\ &\quad= \begin{pmatrix} X^2_{s,u} & \pi_{12}(Z)_{s,u} \\ \pi_{21}(Z)_{s,u} & Y^2_{s,u} \end{pmatrix}\end{aligned}$$

so that

$$X^2_{s,t} + X^2_{t,u} + X^1_{s,t} \otimes X^1_{t,u} = X^2_{s,u},$$
$$\pi_{1,2}(Z)_{s,t} + \pi_{1,2}(Z)_{t,u} + X^1_{s,t} \otimes Y^1_{t,u} = \pi_{1,2}(Z)_{s,u},$$
$$\pi_{2,1}(Z)_{s,t} + \pi_{2,1}(Z)_{t,u} + Y^1_{s,t} \otimes X^1_{t,u} = \pi_{1,2}(Z)_{s,u},$$
$$Y^2_{s,t} + Y^2_{t,u} + Y^1_{s,t} \otimes Y^1_{t,u} = Y^2_{s,u}$$

which implies

$$\left(V_{Z,2}(\varepsilon)_{s,t} \otimes V_{Z,2}(\varepsilon)_{t,u}\right)^2 = V_{Z,2}(\varepsilon)^2_{s,u}. \qquad \square$$

Remark 10. The path $\varepsilon \to V_Z(\varepsilon)$ defined above is continuous in p variation topology.

Now, we can use the above to give a partial description of the tangent space by realizing a representation of tangents to curves of rough paths which arise as of lifts of finite variation paths. Suppose that x and y are finite variation V valued paths. If, for each ε, $C(\varepsilon)$ is the canonical lift of the path $x + \varepsilon y$, then

$$C'(0) = \left(0, C'(0)^1, C'(0)^2\right)$$

where we can express the second term as a sum of iterated integrals

$$C'(0)^2 = \int dx \otimes dC'(0)^1 + \int dC'(0)^1 \otimes dx.$$

The above iterated integrals are also the projections $\pi_{1,2}(Z)$ and $\pi_{2,1}(Z)$ of a rough path $Z \in \Omega_p(V \oplus V)$ which is the canonical lift of $(x, C'(0)^1)$. Hence, we can associate Z to the equivalence class of which $C(\varepsilon)$ is a representative through the relation $\frac{d}{d\varepsilon}|_{\varepsilon=0} V_Z(\varepsilon) = \frac{d}{d\varepsilon}|_{\varepsilon=0} C(\varepsilon)$. Then, in order to have a unique association, we must quotient out $\pi_2(Z)^2$.

It is easy to guess that for a general curve of rough paths $C(\varepsilon)$ starting at X, one should have a relation similar to

$$C'(0)^2 = \pi_{1,2}(Z) + \pi_{2,1}(Z)$$

for Z such that $\pi_1(Z) = X$ and $\pi_2(Z)^1 = C'(0)^1$. The reason why this is not the complete description in the case of a general curve of rough paths is that the second order term in the lift of a finite variation path is completely determined by the first order terms but in general, one may perturb both the first and second order terms independently. Hence, we need something further which is given by the independent second level variation of the rough path X. Since the sum $\pi_{1,2}(Z) + \pi_{2,1}(Z)$ corresponds to second level variation induced by the first level and we will introduce a term φ which is, roughly speaking, the difference between the second level variation induced by the first level and the independent second level variation.

If we take independent second level variation into account, we can give a description of the tangent space using the Lyons–Victoir extension. Suppose $X = (1, X^1, X^2)$ is a p-geometric rough path over V and $C(\varepsilon)$ is a curve in $G\Omega_p(\mathbb{R}^d)$ such that $\frac{d}{d\varepsilon} C(\varepsilon)^1$ has finite p variation and $C(0) = X$. We can form the Lyons–Victoir extension $W \in WG\Omega_p(V \oplus V)$ of $W^1 = (X^1, c'(0)^1)$. Intuitively, the projections $\pi_{1,2}(W)$ and $\pi_{2,1}(W)$ are some cross iterated integrals of X and $C'(0)^1$ and their sum can be thought of as the second level variation of X induced by the first level. However, the sum of the projections $\pi_{1,2}$ and $\pi_{2,1}$ may not equal $C'(0)^2$ due to independent second level variation. Therefore, we introduce $\varphi = C'(0)^2 - \pi_{1,2}(W) - \pi_{2,1}(W)$ which measures the relationship of the extension of $(X^1, C'(0)^1)$ to the derivative. We then form a new element Z of $WG\Omega_p(V \oplus V)$ by replacing $\pi_1(W)^2$ with X^2. The pair (Z, φ) then has all the information contained in the derivative of $C(\varepsilon)$, though we need to quotient out some information because $\pi_2(Z)^2$ is not related to the derivative of $C(\varepsilon)$.

We now formalize the above considerations by first describing the equivalence relation on the pairs (Z, φ).

Definition 11. For $Z, \tilde{Z} \in WG\Omega_p(V \oplus V)$ such that $\pi_1(Z) = \pi_1(\tilde{Z}) = X$, and $\varphi, \tilde{\varphi} \in \Omega_{\frac{p}{2}}(V \otimes V)$, we say the pair (Z, φ) is equivalent to $(\tilde{Z}, \tilde{\varphi})$, denoted $(Z, \varphi) \sim (\tilde{Z}, \tilde{\varphi})$, if

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V_{(Z,\varphi)} = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V_{(\tilde{Z},\tilde{\varphi})}$$

where the variational curve $V^1_{(Z,\varphi)}(\varepsilon) \in C_0(\Delta_T; T^2(V))$ is defined by

$$V^1_{(Z,\varphi)}(\varepsilon) = X^1 + \varepsilon \pi_2(Z)^1,$$
$$V^2_{(Z,\varphi)}(\varepsilon) = X^2 + \varepsilon\big[\pi_{1,2}(Z) + \pi_{2,1}(Z) + \varphi\big] + \varepsilon^2 \pi_2(Z)^2.$$

Remark 12. By Proposition 9, $V_{(Z,\varphi)}$ is actually in $WG\Omega_p(V)$.

Proposition 13. *The above relation $\sim$ is an equivalence relation.*

Proof. This is clear because the relation is defined by an equality. $\square$

We now show that the derivative of a curve of rough paths is also the derivative of a variational curve. Therefore equivalence classes of curves with equal derivative can be uniquely associated to equivalence classes of variational curves which are associated to equivalence classes of pairs (Z, φ). This will justify the following definition.

Definition 14. The collection of all equivalence classes $[Z, \varphi]$ with $Z \in WG\Omega_p(V \oplus V)$ satisfying $\pi_1(Z) = X$ and $\varphi \in WG\Omega_{\frac{p}{2}}(V \otimes V)$ is called the tangent space at X and is denoted by $T_X WG\Omega_p$. Each individual equivalence class is called a tangent vector at X.

Theorem 15. *Let $C(\varepsilon) : [-\tau, \tau] \to WG\Omega_p(V)$ be such that $C(0) = X$ and let*

$$\left(0, C'(0)^1, C'(0)^2\right)$$

be its derivative at 0. If $C'(0)^1$ has finite p-variation and $C'(0)^2$ has finite $\frac{p}{2}$-variation, there exists a unique tangent vector at X $[Z, \varphi]$ such that

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} C(\varepsilon) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} V_{[Z,\varphi]}.$$

Proof. If W is a chosen Lyons–Victoir extension of $(X^1, C'(0)^1)$ and

$$\varphi_{s,t} := C'(0)^2_{s,t} - \pi_{1,2}(W)_{s,t} - \pi_{2,1}(W)_{s,t},$$

form Z defined by

$$Z^1 = \left(X^1, C'(0)^1\right),$$
$$Z^2 = \begin{pmatrix} X^2 & \pi_{1,2}(W) \\ \pi_{2,1}(W) & \pi_2(W)^2 \end{pmatrix}.$$

Such Z is multiplicative, since X and W are. Let us verify that φ is additive now. Since the curve $C(\varepsilon)$ is multiplicative for each ε, we have

$$C(\varepsilon)^2_{s,u} = C(\varepsilon)^2_{s,t} + C(\varepsilon)^2_{t,u} + C(\varepsilon)^1_{s,t} \otimes C(\varepsilon)^1_{t,u}$$

which implies

$$C'(0)^2_{s,u} - C'(0)^2_{s,t} - C'(0)^2_{t,u} = C'(0)^1_{s,t} \otimes X^1_{t,u} + X^1_{s,t} \otimes C'(0)^1_{t,u}$$

but also, by the construction of W, we have

$$\pi_{1,2}(W)_{s,u} - \pi_{1,2}(W)_{s,t} - \pi_{1,2}(W)_{t,u} = X^1_{s,t} \otimes C'(0)^1_{t,u}$$

and

$$\pi_{2,1}(W)_{s,u} - \pi_{2,1}(W)_{s,t} - \pi_{2,1}(W)_{t,u} = C'(0)^1_{s,t} \otimes X^1_{t,u}.$$

Hence,

$$C'(0)^2_{s,u} - \pi_{1,2}(W)_{s,u} - \pi_{2,1}(W)_{s,u} = C'(0)^2_{s,t} - \pi_{1,2}(W)_{s,t} - \pi_{2,1}(W)_{s,t} \\ + C'(0)^2_{t,u} - \pi_{1,2}(W)_{t,u} - \pi_{2,1}(W)_{t,u}$$

which shows that $\varphi_{s,t} = C'(0)^2_{s,t} - \pi_{1,2}(W)_{s,t} - \pi_{2,1}(W)_{s,t}$ is additive. Then, by definition the associated variational curve $V_{[Z,\varphi]}(\varepsilon) \in WG\Omega_p(V)$ where

$$V_{[Z,\varphi]}(\varepsilon)^1 = X^1 + \varepsilon C'(0)^1,$$
$$V_{[Z,\varphi]}(\varepsilon)^2 = X^2 + \varepsilon\big(\pi_{1,2}(W) + \pi_{2,1}(W) + \varphi\big) + \varepsilon^2 \pi_2(W)^2.$$

And by construction $\frac{d}{d\varepsilon}|_{\varepsilon=0} V_{[Z,\varphi]} = \frac{d}{d\varepsilon}|_{\varepsilon=0} C(\varepsilon)$. $\square$

Using the above theorem, we can also define tangents of a curve with the following definition. It will be use in the next section to interpret the flow equation.

Definition 16. If $C(\varepsilon) : [-\sigma, \sigma] \to WG\Omega_p(V)$ is a curve of rough paths, then the tangent of C at τ, denoted $\dot{C}(\tau)$, is the tangent vector $[Z,\varphi](\tau) \in T_{C(\tau)} WG\Omega_p$ satisfying $\pi_1(Z) = C(\tau)$ together with

$$\pi_2(Z)^1 = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} C(\tau + \epsilon)^1$$

and

$$\varphi = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} C(\tau + \epsilon)^2 - \pi_{1,2}(Z) - \pi_{2,1}(Z).$$

Let us show that $T_X WG\Omega_p$ is a linear space.

Lemma 17. *Given $Z_1 = (X, Y)$, $Z_2 = (X, \tilde{Y}) \in G\Omega_p(V \oplus W)$ and $W \in WG\Omega_p(W)$, a particular choice of Lyons–Victoir extension of $(1, Y^1 + \tilde{Y}^1)$, $Z \in C_0(\Delta_T; T^2(V \oplus W))$ defined by*

$$Z^1_{s,t} = \big(X^1_{s,t}, Y^1_{s,t} + \tilde{Y}^1_{s,t}\big),$$
$$Z^2_{s,t} = \begin{pmatrix} X^2_{s,t} & \sum_{i=1}^2 \pi_{1,2}(Z_i)_{s,t} \\ \sum_{i=1}^2 \pi_{2,1}(Z_i)_{s,t} & W^2_{s,t} \end{pmatrix}$$

is a weakly geometric p-rough path over $V \oplus W$.

Proof. Consider $(1, Y^1 + \tilde{Y}^1)$. We can extend this via the Lyons–Victoir extension to a weakly geometric rough path $W \in WG\Omega_p(W)$ which we write as

$$W = \big(1, Y^1 + \tilde{Y}^1, W^2\big).$$

By definition,

$$(Z_{s,t} \otimes Z_{t,u})^2_{1,1} = X^2_{s,t} + X^2_{t,u} + X^1_{s,t} \otimes X^1_{t,u},$$
$$(Z_{s,t} \otimes Z_{t,u})^2_{1,2} = \sum_{i=1}^2 \pi_{1,2}(Z_i)_{s,t} + \sum_{i=1}^2 \pi_{1,2}(Z_i)_{t,u} + \big(X^1_{s,t} \otimes Y^1_{t,u} + X^1_{s,t} \otimes \tilde{Y}^1_{t,u}\big),$$
$$(Z_{s,t} \otimes Z_{t,u})^2_{2,1} = \sum_{i=1}^2 \pi_{2,1}(Z_i)_{s,t} + \sum_{i=1}^2 \pi_{2,1}(Z_i)_{t,u} + \big(Y^1_{s,t} \otimes X^1_{t,u} + \tilde{Y}^1_{s,t} \otimes X^1_{t,u}\big),$$
$$(Z_{s,t} \otimes Z_{t,u})^2_{2,2} = W^2_{s,t} + W^2_{t,u} + Y^1_{s,t} \otimes Y^1_{t,u} + Y^1_{s,t} \otimes \tilde{Y}^1_{t,u} + \tilde{Y}^1_{s,t} \otimes Y^1_{t,u} + \tilde{Y}^1_{s,t} \otimes \tilde{Y}^1_{t,u}.$$

Since Z_i is multiplicative for $i=1,2$ we have

$$\begin{pmatrix} X^2_{s,t}+X^2_{t,u}+X^1_{s,t}\otimes X^1_{t,u} & \pi_{1,2}(Z_1)_{s,t}+\pi_{1,2}(Z_1)_{t,u}+X^1_{s,t}\otimes Y^1_{t,u} \\ \pi_{2,1}(Z_1)_{s,t}+\pi_{2,1}(Z_1)_{t,u}+Y^1_{s,t}\otimes X^1_{t,u} & Y^2_{s,t}+Y^2_{t,u}+Y^1_{s,t}\otimes Y^1_{t,u} \end{pmatrix}$$
$$= \begin{pmatrix} X^2_{s,u} & \pi_{1,2}(Z_1)_{s,u} \\ \pi_{2,1}(Z_1)_{s,u} & Y^2_{s,u} \end{pmatrix}$$

and

$$\begin{pmatrix} X^2_{s,t}+X^2_{t,u}+X^1_{s,t}\otimes X^1_{t,u} & \pi_{1,2}(Z_2)_{s,t}+\pi_{1,2}(Z_2)_{t,u}+X^1_{s,t}\otimes \tilde{Y}^1_{t,u} \\ \pi_{2,1}(Z_2)_{s,t}+\pi_{2,1}(Z_2)_{t,u}+\tilde{Y}^1_{s,t}\otimes X^1_{t,u} & \tilde{Y}^2_{s,t}+\tilde{Y}^2_{t,u}+\tilde{Y}^1_{s,t}\otimes \tilde{Y}^1_{t,u} \end{pmatrix}$$
$$= \begin{pmatrix} X^2_{s,u} & \pi_{1,2}(Z_2)_{s,u} \\ \pi_{2,1}(Z_2)_{s,u} & \tilde{Y}^2_{s,u} \end{pmatrix}.$$

Also, since W is multiplicative,

$$W^2_{s,t}+W^2_{t,u}+\left(Y^1+\tilde{Y}^1\right)\otimes\left(Y^1+\tilde{Y}^1\right)=W_{s,u}.$$

Putting these expressions together, we get $(Z_{s,t}\otimes Z_{t,u})^2=Z^2_{s,u}$. $\square$

Proposition 18. *The tangent space $T_X G\Omega_p(V)$ is linear. More precisely the operations*

$$[Z_1,\varphi_1]+[Z_2,\varphi_2]:=[Z,\varphi_1+\varphi_2],$$

where

$$Z^1=\left(X^1_{s,t},\pi_2(Z_1)^1+\pi_2(Z_2)^1\right),$$
$$Z^2=\begin{pmatrix} X^2_{s,t} & \sum_{i=1}^2 \pi_{1,2}(Z_i)_{s,t} \\ \sum_{i=1}^2 \pi_{2,1}(Z_i)_{s,t} & W^2_{s,t} \end{pmatrix}$$

for W some Lyons–Victoir extension of $(1,\pi_2(Z_1)^1+\pi_2(Z_2)^1)$*, and*

$$\lambda[Z,\varphi]:=[Z_\lambda,\lambda\varphi]$$

where Z_λ is defined by

$$Z^1_\lambda=\left(X^1_{s,t},\lambda\pi_2(Z)^1\right),$$
$$Z^2_\lambda=\begin{pmatrix} X^2_{s,t} & \lambda\pi_{1,2}(Z)_{s,t} \\ \lambda\pi_{2,1}(Z)_{s,t} & \lambda^2 Z^2_{s,t} \end{pmatrix}$$

are well defined.

Proof. That Z, as defined above, is in $WG\Omega_p(V\oplus V)$ follows from Lemma 17 so $[Z,\varphi_1+\varphi_2]$ is indeed an equivalence class. We now show that $[Z,\varphi_1+\varphi_2]$ does not depend on the choice of representatives from $[Z_1,\varphi_1]$ and $[Z_2,\varphi_2]$. Let $(\tilde{Z}_1,\tilde{\varphi}_1)$ and $(\tilde{Z}_2,\tilde{\varphi}_2)$ be some other representatives from the equivalence classes $[Z_1,\varphi_1]$ and $[Z_2,\varphi_2]$ respectively. We form $(\tilde{Z},\tilde{\varphi}_1+\tilde{\varphi}_2)$ with

$$\tilde{Z}^1=\left(X^1_{s,t},\pi_2(\tilde{Z}_1)^1+\pi_2(\tilde{Z}_2)^1\right),$$
$$\tilde{Z}^2=\begin{pmatrix} X^2_{s,t} & \sum_{i=1}^2 \pi_{1,2}(\tilde{Z}_i)_{s,t} \\ \sum_{i=1}^2 \pi_{2,1}(\tilde{Z}_i)_{s,t} & W^2_{s,t} \end{pmatrix}.$$

Then

$$V^1_{(\tilde{Z},\tilde{\varphi}_1+\tilde{\varphi}_2)}(\varepsilon) = X^1 + \varepsilon\big[\pi_2(\tilde{Z}_1)^1 + \pi_2(\tilde{Z}_2)^1\big],$$

$$V^2_{(\tilde{Z},\tilde{\varphi}_1+\tilde{\varphi}_2)}(\varepsilon) = X^2 + \varepsilon\sum_{i=1}^{2}\pi_{1,2}(\tilde{Z}_i)_{s,t} + \varepsilon\sum_{i=1}^{2}\pi_{2,1}(\tilde{Z}_i)_{s,t} + \varepsilon(\tilde{\varphi}_1 + \tilde{\varphi}_2) + \varepsilon^2 W^2_{s,t}.$$

Furthermore, since $(Z_i, \varphi_i) \sim (\tilde{Z}_i, \tilde{\varphi}_i)$ for $i \in \{1, 2\}$, we have

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V_{(Z_i,\varphi_i)}(\varepsilon) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V_{(\tilde{Z}_i,\tilde{\varphi}_i)}(\varepsilon)$$

which gives

$$\pi_2(Z_i)^1 = \pi_2(\tilde{Z}_i)^1$$

and

$$\pi_{1,2}(Z_i) + \pi_{2,1}(Z_i) + \varphi_i = \pi_{1,2}(\tilde{Z}_i) + \pi_{2,1}(\tilde{Z}_i) + \tilde{\varphi}_i.$$

Therefore,

$$\begin{aligned}
\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V^1_{(\tilde{Z},\tilde{\varphi}_1+\tilde{\varphi}_2)}(\varepsilon) &= \big[\pi_2(\tilde{Z}_1)^1 + \pi_2(\tilde{Z}_2)^1\big] \\
&= \big[\pi_2(Z_1)^1 + \pi_2(Z_2)^1\big] \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V^1_{(Z,\varphi_1+\varphi_2)}(\varepsilon)
\end{aligned}$$

and also

$$\begin{aligned}
\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V^2_{(\tilde{Z},\tilde{\varphi}_1+\tilde{\varphi}_2)}(\varepsilon) &= \sum_{i=1}^{2}\pi_{1,2}(\tilde{Z}_i)_{s,t} + \sum_{i=1}^{2}\pi_{2,1}(\tilde{Z}_i)_{s,t} + \tilde{\varphi}_1 + \tilde{\varphi}_2 \\
&= \sum_{i=1}^{2}\pi_{1,2}(Z_i)_{s,t} + \sum_{i=1}^{2}\pi_{2,1}(Z_i)_{s,t} + \varphi_1 + \varphi_2 \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} V^2_{(Z,\varphi_1+\varphi_2)}(\varepsilon).
\end{aligned}$$

Hence $(\tilde{Z}, \tilde{\varphi}_1 + \tilde{\varphi}_2) \sim (Z, \varphi_1 + \varphi_2)$ and the addition is well defined. For scalar multiplication, $Z_\lambda \in WG\Omega_p(V)$ follows from Proposition 9. The fact that it is independent of the choice of equivalence class representative is the same in nature as the proof of this fact for addition. □

4. Flow equations

In this section we investigate the existence and uniqueness of solutions to flow equations on the space of weakly geometric rough paths $WG\Omega_p$ given by

$$\dot{C}(\tau) = F\big(C(\tau)\big), \qquad C(0) = X.$$

We stress that here the term $\dot{C}(\tau)$ is not the derivative in the sense of Definition 8 but is the tangent to the curve at τ in the sense of Definition 16 while F assigns elements of the tangent space at $C(\tau)$ to each $C(\tau)$. Key to the constructions in this section is that although the collection of tangent spaces does not have a vector bundle structure due to lack of local trivialization, there is some useful structure which comes from rough paths being imbedded in the linear space $C(\Delta; T^2(V))$ equipped with the p-variation distance.

Throughout this section, we consider a class of vector fields given as follows. Suppose we are given two maps $Z : WG\Omega_p(V) \to WG\Omega_p(V \oplus V)$ and $\varphi : WG\Omega_p(V) \to WG\Omega_{\frac{p}{2}}(V)$ such that for each X, $\pi_1(Z(X)) = X$. Then, for each X, we can form the tangent vector $[Z(X), \varphi(X)]$ and so obtain an associated vector field $F(X)$. We will use the notation $F_{Z,\varphi}(X) := \pi_{1,2}(Z(X)) + \pi_{2,1}(Z(X)) + \varphi(X)$ together with $F_Z(X) = Z(X)$ and $F_\varphi(X) = \varphi(X)$ when we wish to emphasize the tangent field generated by Z and φ.

First let us define the following function on the tangent collection for future clarity of the notation.

Definition 19. For $[Z,\varphi]_X \in T_X WG\Omega_p(V)$ and $[W,\phi]_Y \in T_Y WG\Omega_p(V)$ let

$$\begin{aligned}\tilde{d}_p\big([Z,\varphi],[W,\phi]\big) := \max\Bigg[& d_p\big(\pi_1(Z),\pi_1(W)\big),\\ & \sup_{\mathcal{D}}\bigg\{\sum_l \big|\pi_2(Z)^1_{t_l,t_{l+1}} - \pi_2(W)^1_{t_l,t_{l+1}}\big|^p\bigg\}^{\frac{1}{p}},\\ & \sup_{\mathcal{D}}\bigg\{\sum_l \big|\big[\pi_{1,2}(Z) + \pi_{2,1}(Z) + \varphi\big]_{t_l,t_{l+1}}\\ & - \big[\pi_{1,2}(W) + \pi_{2,1}(W) + \phi\big]_{t_l,t_{l+1}}\big|^{\frac{p}{2}}\bigg\}^{\frac{2}{p}}\Bigg]\end{aligned}$$

which does not depend on the choice of representative of the equivalence class.

Proposition 20. *The function $\tilde{d}_p$ is a metric on the disjoint union of the tangent spaces, $TWG\Omega_p(V)$.*

Proof. Non-negativity and symmetry are inherited from the properties of the norm on V. If $\tilde{d}_p([Z,\varphi]_X,[W,\phi]_Y) = 0$ then $\pi_1(Z) = \pi_1(W)$ so that $[Z,\varphi]$ and $[W,\phi]$ are in the same tangent space. By Definition 11, $\pi_2(Z)^1 = \pi_2(W)^1$ and

$$\pi_{1,2}(Z) + \pi_{2,1}(Z) + \varphi = \pi_{1,2}(W) + \pi_{2,1}(W) + \phi$$

imply that they are in the same equivalence class. Hence $\tilde{d}_p([Z,\varphi],[W,\phi]) = 0$ if and only if $[Z,\varphi]_X = [W,\phi]_Y$.

For the triangle inequality, consider $[S,\sigma] \in T_R WG\Omega_p(V)$. The triangle inequality for d_p and the Minkowski inequality give

$$d_p\big(\pi_1(Z),\pi_1(W)\big) \leqslant d_p\big(\pi_1(Z),\pi_1(S)\big) + d_p\big(\pi_1(S),\pi_1(W)\big),$$

$$\begin{aligned}\bigg\{\sum_l \big|\pi_2(Z)^1_{t_l,t_{l+1}} - \pi_2(W)^1_{t_l,t_{l+1}}\big|^p\bigg\}^{\frac{1}{p}} \leqslant{}& \bigg\{\sum_l \big|\pi_2(Z)^1_{t_l,t_{l+1}} - \pi_2(S)^1_{t_l,t_{l+1}}\big|^p\bigg\}^{\frac{1}{p}}\\ &+ \bigg\{\sum_l \big|\pi_2(S)^1_{t_l,t_{l+1}} - \pi_2(W)^1_{t_l,t_{l+1}}\big|^p\bigg\}^{\frac{1}{p}},\end{aligned}$$

and

$$\begin{aligned}
&\left\{\sum_l \left|\left[\pi_{1,2}(Z)+\pi_{2,1}(Z)+\varphi\right]_{t_l,t_{l+1}} - \left[\pi_{1,2}(W)+\pi_{2,1}(W)+\phi\right]_{t_l,t_{l+1}}\right|^{\frac{p}{2}}\right\}^{\frac{2}{p}} \\
&\quad \leqslant \left\{\sum_l \left|\left[\pi_{1,2}(Z)+\pi_{2,1}(Z)+\varphi\right]_{t_l,t_{l+1}} - \left[\pi_{1,2}(S)+\pi_{2,1}(S)+\sigma\right]_{t_l,t_{l+1}}\right|^{\frac{p}{2}}\right\}^{\frac{2}{p}} \\
&\qquad + \left\{\sum_l \left|\left[\pi_{1,2}(S)+\pi_{2,1}(S)+\sigma\right]_{t_l,t_{l+1}} - \left[\pi_{1,2}(W)+\pi_{2,1}(W)+\phi\right]_{t_l,t_{l+1}}\right|^{\frac{p}{2}}\right\}^{\frac{2}{p}}.
\end{aligned}$$

Therefore, by definition of the supremum as the least upper bound, in the previous two equations the supremum over all partitions on the left-hand side is less than the sum of the supremums on the right-hand side. Finally, bounding each of

$$d_p\big(\pi_1(Z),\pi_1(W)\big), \qquad \sup_{\mathcal{D}}\left\{\sum_l \left|\pi_2(Z)^1_{t_l,t_{l+1}} - \pi_2(W)^1_{t_l,t_{l+1}}\right|^p\right\}^{\frac{1}{p}},$$

and

$$\sup_{\mathcal{D}}\left\{\sum_l \left|\left[\pi_{1,2}(Z)+\pi_{2,1}(Z)+\varphi\right]_{t_l,t_{l+1}} - \left[\pi_{1,2}(W)+\pi_{2,1}(W)+\phi\right]_{t_l,t_{l+1}}\right|^{\frac{p}{2}}\right\}^{\frac{2}{p}}$$

by the sum

$$\begin{aligned}
&\max\Bigg[d_p\big(\pi_1(Z),\pi_1(S)\big), \sup_{\mathcal{D}}\left\{\sum_l \left|\pi_2(Z)^1_{t_l,t_{l+1}} - \pi_2(S)^1_{t_l,t_{l+1}}\right|^p\right\}^{\frac{1}{p}}, \\
&\qquad \sup_{\mathcal{D}}\left\{\sum_l \left|\left[\pi_{1,2}(Z)+\pi_{2,1}(Z)+\varphi\right]_{t_l,t_{l+1}} - \left[\pi_{1,2}(S)+\pi_{2,1}(S)+\sigma\right]_{t_l,t_{l+1}}\right|^{\frac{p}{2}}\right\}^{\frac{2}{p}}\Bigg] \\
&\quad + \max\Bigg[d_p\big(\pi_1(S),\pi_1(W)\big), \sup_{\mathcal{D}}\left\{\sum_l \left|\pi_2(S)^1_{t_l,t_{l+1}} - \pi_2(W)^1_{t_l,t_{l+1}}\right|^p\right\}^{\frac{1}{p}}, \\
&\qquad \sup_{\mathcal{D}}\left\{\sum_l \left|\left[\pi_{1,2}(S)+\pi_{2,1}(S)+\sigma\right]_{t_l,t_{l+1}} - \left[\pi_{1,2}(W)+\pi_{2,1}(W)+\phi\right]_{t_l,t_{l+1}}\right|^{\frac{p}{2}}\right\}^{\frac{2}{p}}\Bigg]
\end{aligned}$$

gives the result. □

Now we are able to state our Lipschitz condition on F which allows existence solutions.

Definition 21. Let $Z : WG\Omega_p(V) \to WG\Omega_p(V \oplus V)$ and $\varphi : WG\Omega_p(V) \to WG\Omega_{\frac{p}{2}}(V)$ such that $\pi_1(Z(X)) = X$ and form the vector field F such that $F(X) = [Z(X), \varphi(X)]$.

We will say such an F is locally Lipschitz near X_0 if there exist positive constants r, C_1, and C_2 such that for all $X, Y \in B_r(X_0)$

$$\begin{aligned}
&\tilde{d}_p\big(F(X), F(Y)\big) \leqslant C_1 d_p(X,Y), \\
&\tilde{d}_q\big(F(X), F(Y)\big) \leqslant C_2 d_p(X,Y)
\end{aligned}$$

for $q > p$.

We will say F is globally Lipschitz if the above relations hold for all $X, Y \in WG\Omega_p(V)$.

Now let us introduce our concept of a solution to the flow equation for a non-differentiable curve. We know from the definition of the equivalence class, that two tangents (at a fixed point X) are in the same equivalence class if the derivatives at $\varepsilon = 0$ of their variational curves are equal, i.e. $(Z, \varphi) \sim (\tilde{Z}, \tilde{\varphi})$ if and only if $\frac{d}{d\varepsilon}|_{\varepsilon=0} V_{(Z,\varphi)}(\varepsilon) = \frac{d}{d\varepsilon}|_{\varepsilon=0} V_{(\tilde{Z},\tilde{\varphi})}(\varepsilon)$. Also, using Theorem 15, we can associate any differentiable curve U to a unique equivalence class $[Z, \varphi]$ such that $\frac{d}{d\varepsilon}|_{\varepsilon=0} U(\varepsilon) = \frac{d}{d\varepsilon}|_{\varepsilon=0} V_{[Z,\varphi]}(\varepsilon)$. Hence, a curve $U(\tau)$ differentiable in our sense is a classical solution to the flow equation if

$$\lim_{\varepsilon \to 0} \frac{U(\tau + \varepsilon) - U(\tau)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{V_{[F_Z(U(\tau)), F_\varphi(U(\tau))]}(\varepsilon) - U(\tau)}{\varepsilon}$$

for each $\tau \in [0, T]$. Formally rearranging this equation leaves us with

$$\lim_{\varepsilon \to 0} \frac{U(\tau + \varepsilon) - V_{[F_Z(U(\tau)), F_\varphi(U(\tau))]}(\varepsilon)}{\varepsilon} = 0.$$

This second expression can be defined even for some non-differentiable curves. The natural metric in which this limit should take place is of course the p-variation metric. However, the following definition uses q (for $q > p$) instead of p due to technical limitations which will become clear in the proof.

Definition 22. A continuous curve $U : [0, T] \to \Omega_q(V)$ is said to be a solution to the flow equation

$$\dot{C}(\tau) = F\big(C(\tau)\big),$$
$$C(0) = X$$

if $U(0) = X$ and

$$\lim_{h \downarrow 0} h^{-1}\big[d_q\big(U(\tau + h), V_{[F_Z(U(\tau)), F_\varphi(U(\tau))]}(h)\big)\big] = 0$$

for each $\tau \in [0, T]$ and some $q > p$.

U is said to be an ε-solution to the flow equation if $U(0) = X$ and

$$\lim_{h \downarrow 0} h^{-1}\big[d_p\big(U(\tau + h), V_{[F_Z(U(\tau)), F_\varphi(U(\tau))]}(h)\big)\big] \leqslant \varepsilon$$

for each $\tau \in [0, T]$.

Roughly speaking, an ε-solution is a curve whose tangent at a point is in some sense close to the tangent assigned by F at that point. Note that if we were in a Banach space case then the above would be

$$\lim_{h \downarrow 0} \left| \frac{U(\tau + h) - [U(\tau) + hF(U(\tau))]}{h} \right|,$$

i.e. it would be equivalent to the statement: $U'(\tau)$ exists and equals $F(U(\tau))$.

Remark 23. Recall that for a given tangent vector, a variational curve is not unique. In fact there are infinitely many choices each associated to a different Lyons–Victoir extension, indeed a variational curve for a tangent at X has component in $V \otimes V$ given by $V_{[Z,\varphi]}(\varepsilon)^2 = X^2 + \varepsilon(\pi_{1,2}(Z) + \pi_{2,1}(Z) + \varphi) + \varepsilon^2(W^2)$ where W is an extension of $X^1 + \pi_2(Z)^1$. Therefore, the above definition may depend on the choice of variational curve which we associate to the tangent $[F_Z(U(\tau)), F_\varphi(U(\tau))]$. However, for the class of vector fields we consider, the canonical choice of W^2 is given by $\pi_2(Z(U(\tau)))^2$. We will always assume we make this choice and therefore, we refer to the variational curve rather than a variational curve.

In order to obtain solutions, we first construct an approximate solution given by an Euler approximation scheme. The only issue is correctly choosing the Euler "polygon" approximation through the variational curve associated to a tangent vector.

Lemma 24. *Let F be a locally Lipschitz vector field near the rough path X_0 in the sense of Definition* 21*, define*

$$M = \sup_{X \in B_r(X_0)} \left[\tilde{d}_p\big(F(X), 0\big), \sup_{\mathcal{D}} \left\{ \sum_l \left|\pi_2\big(Z(X)\big)^2_{t_l, t_{l+1}}\right|^{\frac{p}{2}} \right\}^{\frac{2}{p}} \right],$$

and set $\alpha = \frac{r}{2M}$. Then for all $\varepsilon > 0$ there exists an ε-solution $U \in C([0,\alpha] : B_r(X_0))$ of the flow equation with initial value $X_0 \in WG\Omega_p(V)$ satisfying

$$d_p(U_\tau, U_\sigma) \leqslant (1 + 2\alpha) M |\tau - \sigma|, \quad \forall \sigma, \tau \in [0, \alpha]. \tag{4.1}$$

Proof. Let $\varepsilon > 0$ be given and partition the interval $[0, \alpha]$ into

$$0 = \tau_0 < \tau_1 < \cdots < \tau_n = \alpha$$

such that

$$\max_i |\tau_{i+1} - \tau_i| \leqslant \max\left\{ \frac{\varepsilon}{(C_1 + 2)M}, 1 \right\}$$

where C_1 is the p-variation Lipschitz constant of F as expressed in Definition 21. For $\tau \in [\tau_i, \tau_{i+1})$, define inductively

$$U(\tau) = V_{[F(U_{\tau_i})]}(\tau - \tau_i)$$

where $V_{[F(U_{\tau_i})]}$ is the variational curve associated to the tangent $F(U_{\tau_i})$. More precisely, we first define the points

$$\begin{aligned} U^1_{\tau_1} &= X^1_0 + \tau_1 \pi_2\big(F(X_0)\big)^1, \\ U^2_{\tau_1} &= X^2_0 + \tau_1 \big[\pi_{1,2}\big(F(X_0)\big) + \pi_{2,1}\big(F(X_0)\big) + F_\varphi(X_0)\big] + + \tau_1^2 \pi_2\big(Z(X_0)\big)^2 \end{aligned}$$

and

$$\begin{aligned} U^1_{\tau_{i+1}} &= U^1_{\tau_i} + (\tau_{i+1} - \tau_i) \pi_2\big(F(U_{\tau_i})\big)^1, \\ U^2_{\tau_{i+1}} &= U^2_{\tau_i} + (\tau_{i+1} - \tau_i) \big[\pi_{1,2}\big(F(U_{\tau_i})\big) + \pi_{2,1}\big(F(U_{\tau_i})\big) + F_\varphi(U_{\tau_i})\big] \\ &\quad + (\tau_{i+1} - \tau_i)^2 \pi_2\big(Z(U_{\tau_i})\big)^2 \end{aligned}$$

for $i \in \{1, \dots, n\}$ and then the curve is defined by

$$\big(V_{[F(U_{\tau_i})]}(\tau)\big)^1 = U^1_{\tau_i} + (\tau-\tau_i)\pi_2\big(F(U_{\tau_i})\big)^1,$$
$$\big(V_{[F(U_{\tau_i})]}(\tau)\big)^2 = U^2_{\tau_i} + (\tau-\tau_i)\big[\pi_{1,2}\big(F(U_{\tau_i})\big)+\pi_{2,1}\big(F(U_{\tau_i})\big)+F_\varphi(U_{\tau_i})\big] + (\tau-\tau_i)^2\pi_2\big(Z(U_{\tau_i})\big)^2$$

whenever $\tau\in[\tau_i,\tau_{i+1})$.

The curve thus defined satisfies $U(\tau)\in B_r(X_0)$. In fact, for $\tau\in[0,\tau_1)$

$$\big\|U^1(\tau)-X^1_0\big\| = \tau\big\|\pi_2\big(F(X_0)\big)^1\big\|,$$
$$\big\|U^2(\tau)-X^2_0\big\| \leqslant \tau\big\|F_{Z,\varphi}(X_0)\big\| + \tau^2\big\|\pi_2\big(F(X_0)\big)^2\big\|$$

so that $d_p(U(\tau),X_0)\leqslant 2\tau_1 M$. Similarly, for $\tau\in[\tau_i,\tau_{i+1})$

$$\big\|U^1(\tau)-U^1(\tau_i)\big\| = (\tau-\tau_i)\big\|\pi_2\big(F\big(U(\tau_i)\big)\big)^1\big\|,$$
$$\big\|U^2(\tau)-U^2(\tau_i)\big\| \leqslant (\tau-\tau_i)\big\|F_{Z,\varphi}\big(U(\tau_i)\big)\big\| + (\tau-\tau_i)^2\big\|\pi_2\big(Z\big(U(\tau_i)\big)\big)^2\big\|$$

implies $d_p(U(\tau_i),U(\tau))\leqslant 2|\tau_{i+1}-\tau_i|M$. Hence, for general $\tau\in[0,\alpha]$,

$$\begin{aligned} d_p\big(U(\tau),X_0\big) &\leqslant d_p(X_0,U_{\tau_1}) + \sum_{i=1}^{k-1} d_p\big(U(\tau_i),U(\tau_{i+1})\big) + d_p\big(U(\tau_k),U(\tau)\big)\\ &\leqslant 2M\tau_1 + 2M\sum_{i=1}^{k-1}|\tau_{i+1}-\tau_i| + 2M|\tau-\tau_k|\\ &= 2M\tau\\ &\leqslant 2M\alpha = r \end{aligned}$$

where k is such that $\tau\in[\tau_k,\tau_{k+1})$ and we have used $(\tau_{i+1}-\tau_i)^2\leqslant(\tau_{i+1}-\tau_i)$ since $(\tau_{i+1}-\tau_i)\leqslant 1$. To see the Lipschitz condition is satisfied, first take $\tau,\sigma\in[\tau_i,\tau_{i+1})$ for which we have

$$U^1_\tau - U^1_\sigma = (\tau-\sigma)\pi_2\big(F_Z(U_{\tau_i})\big)^1,$$
$$U^2_\tau - U^2_\sigma = (\tau-\sigma)F_{Z,\varphi}(U_{\tau_i}) + \big[(\tau-\sigma)+2(\sigma-\tau_i)\big](\tau-\sigma)\pi_2\big(Z(U_{\tau_i})\big)^2$$

so that, by homogeneity of the metric, $d_p(U_\tau,U_\sigma)\leqslant CM|\tau-\sigma|$. For general $\tau,\sigma\in[0,\alpha]$ such that $\sigma\in[\tau_m,\tau_{m+1})$, $\tau\in[\tau_n,\tau_{n+1})$, we have

$$U^1_\tau - U^1_\sigma = \sum_{i=m}^{n}\int_\sigma^\tau \pi_2\big(F(U_{\tau_i})\big)^1\chi_{[\tau_i,\tau_{i+1})}(\gamma)\,d\gamma,$$
$$U^2_\tau - U^2_\sigma = \sum_{i=m}^{n}\int_\sigma^\tau \big[F_{Z,\varphi}(U_{\tau_i}) + 2(\gamma-\tau_i)\pi_2\big(Z(U_{\tau_i})\big)^2\big]\chi_{[\tau_i,\tau_{i+1})}(\gamma)\,d\gamma$$

where we use the Bochner integral in the Banach spaces $(V,\|\cdot\|_V)$ and $(V\otimes V,\|\cdot\|_{V\otimes V})$. Consequently,

$$\big\|\big[U^1_\tau - U^1_\sigma\big]\big\| \leqslant \max_i\big\|\pi_2\big(F(U_{\tau_i})\big)^1\big\||\tau-\sigma|,$$
$$\big\|\big[U^2_\tau - U^2_\sigma\big]\big\| \leqslant \max_i\big[\big\|F_{Z,\varphi}(U_{\tau_i})\big\| + 2\alpha\big\|\pi_2\big(Z(U_{\tau_i})\big)^2\big\|\big]|\tau-\sigma|$$

which implies

$$d_p(U_\tau, U_\sigma) \leqslant (1 + 2\alpha)M|\tau - \sigma|.$$

Now let us show that U is indeed an ε-solution. Whenever $\tau \in [\tau_i, \tau_{i+1})$, for sufficiently small h, we also have $(\tau + h) \in [\tau_i, \tau_{i+1})$ and hence

$$U(\tau + h)^1 = U^1_{\tau_i} + (\tau + h - \tau_i)\pi_2\big(F(U_{\tau_i})\big)^1,$$
$$U(\tau + h)^2 = U^2_{\tau_i} + (\tau + h - \tau_i)F_{Z,\varphi}(U_{\tau_i}) + (\tau + h - \tau_i)^2\pi_2\big(Z(U_{\tau_i})\big)^2$$

and

$$V_{F(U_\tau)}(h)^1 = U^1_\tau + h\pi_2\big(F(U_\tau)\big)^1,$$
$$V_{F(U)_\tau)}(h)^2 = U^2_\tau + hF_{Z,\varphi}(U_\tau) + h^2\pi_2\big(Z(U_\tau)\big)^2.$$

This shows,

$$U(\tau + h)^1 - V_{F(U_\tau)}(h)^1 = \big(U^1_{\tau_i} - U^1_\tau\big) + (\tau - \tau_i)\pi_2\big(F(U_{\tau_i})\big)^1 + h\big(\pi_2\big(F(U_{\tau_i})\big)^1 - \pi_2\big(F(U_\tau)\big)^1\big)$$

and

$$\begin{aligned} &U(\tau + h)^2 - V_{F(U_\tau)}(h)^2 \\ &\quad = \big(U^2_{\tau_i} - U^2_\tau\big) + (\tau - \tau_i)F_{Z,\varphi}(U_{\tau_i}) \\ &\qquad + h\big(F_{Z,\varphi}(U_{\tau_i}) - F_{Z,\varphi}(U_\tau)\big)\big[(\tau - \tau_i)^2 + 2h(\tau - \tau_i)\big]\pi_2\big(Z(U_{\tau_i})\big)^2 \\ &\qquad + h^2\big[\pi_2\big(Z(U_{\tau_i})\big)^2 - \pi_2\big(Z(U_\tau)\big)^2\big]. \end{aligned}$$

Since, $\tau_i < \tau < \tau_{i+1}$

$$U^1_\tau = U^1_{\tau_i} + (\tau - \tau_i)\pi_2\big(F(U_{\tau_i})\big)^1,$$
$$U^2_\tau = U^2_{\tau_i} + (\tau - \tau_i)F_{Z,\varphi}\big(F(U_{\tau_i})\big) + (\tau - \tau_i)^2\pi_2\big(Z(U_{\tau_i})\big)^2$$

and so

$$\big(U^1_{\tau_i} - U^1_\tau\big) + (\tau - \tau_i)\pi_2\big(F(U_{\tau_i})\big)^1 = 0,$$
$$\big(U^2_{\tau_i} - U^2_\tau\big) + (\tau - \tau_i)F_{Z,\varphi}(U_{\tau_i}) + (\tau - \tau_i)^2\pi_2\big(Z(U_{\tau_i})\big)^2 = 0.$$

Therefore,

$$h^{-1}\big(U(\tau + h) - V_{F(U_\tau)}(h)\big)^1 = \big(\pi_2\big(F(U_{\tau_i})\big)^1 - \pi_2\big(F(U_\tau)\big)^1\big), \tag{4.2}$$
$$h^{-1}\big(U(\tau + h) - V_{F(U_\tau)}(h)\big)^2 = \big(F_{Z,\varphi}(U_{\tau_i}) - F_{Z,\varphi}(U_\tau)\big) + 2(\tau - \tau_i)\pi_2\big(Z(U_{\tau_i})\big)^2 + h\big[\pi_2\big(Z(U_{\tau_i})\big)^2 - \pi_2\big(Z(U_\tau)\big)^2\big] \tag{4.3}$$

and by the Lipschitz property of F and homogeneity of the metric,

$$h^{-1}d_p\big(U(\tau + h), V_{F(U_\tau)}(h)\big) \leqslant C_1 d_p(U_{\tau_i}, U_\tau) + 2M|\tau - \tau_i| + 2hM.$$

Also,

$$
\begin{aligned}
&U_\tau^1 = U_{\tau_i}^1 + (\tau - \tau_i)\pi_2\big(F_Z(U_{\tau_i})\big)^1,\\
&U_\tau^2 = U_{\tau_i}^2 + (\tau - \tau_i)F_{Z,\varphi}(U_{\tau_i}),\\
&d_p(U_{\tau_i}, U_\tau) \leqslant M|\tau - \tau_i|.
\end{aligned}
$$

This implies

$$
\begin{aligned}
h^{-1}d_p\big(U(\tau+h), V_{F(U_\tau)}(h)\big) &\leqslant (C_1+2)M|\tau-\tau_i| + 2hM\\
&\leqslant \varepsilon + 2hM
\end{aligned}
$$

so letting $h \to 0$ gives the result. □

Remark 25. If the map Z which defines F also satisfies the Lipschitz property, then the interval on which the ε-solution is defined has length greater than $\frac{1}{C}$ where C depends only on the Lipschitz constants and the initial data. Indeed, for all $X \in B_r(X_0)$, $\tilde{d}_p(F(X), 0) \leqslant C_1 r + C_2$ and $d_p(Z(X), 0) \leqslant C_3 r + C_4$ where the C_i are Lipschitz constants or the distance of X_0 from 0. Hence, $\alpha = \frac{r}{M} \geqslant \frac{r}{C_i r + C_{i+1}}$ and choosing $r \geqslant 1$ gives $\alpha \geqslant \frac{1}{C_i + C_{i+1}}$.

Consequently, we have an approximate solution for any level of closeness on a fixed interval which is independent of how close the approximation is. From here, the Lipschitz property of U given by Eq. (4.1) means that a family consisting of ε_n approximations where $\varepsilon_n \to 0$ as $n \to \infty$ is equicontinuous. Indeed we have the following.

Lemma 26. *Let $\{U^n\}$ be a sequence of paths in $C([0,T], WG\Omega_p)$ such that for each n, U^n is an ε_n-solution to the flow equation where $\varepsilon_n \to 0$ as $n \to \infty$. Then the collection $\{U^n\}$ has a sub-sequence $\{U^{n_k}\}$ which converges in $C([0,T], \Omega_q)$ equipped with uniform topology, for all $q > p$.*

Proof. This follows from the Ascoli–Arzela theorem for metric spaces and Theorem 5. □

A natural question to consider is whether or not the limit of ε_n solutions given by the preceding lemma is a solution. The proof that this is the case relies on the vector field also being Lipschitz in the q variation sense and follows the same arguments as given in Lemma 24.

Lemma 27. *Given a convergent sub-sequence $\{U^n\}$ of ε_n-solutions of the flow equation, the limit $U \in C([0,\alpha]; \Omega_q)$ is a solution of the flow equation.*

Proof. Let us first consider the distance involving the approximating sequence, i.e. $d_q(U^n(\tau+h), V_{F(U^n(\tau))}(h))$. By Eqs. (4.2) and (4.3) in the proof of Lemma 24, we have for $\tau \in [\tau_i, \tau_{i+1})$ and sufficiently small h

$$
\begin{aligned}
h^{-1}\big(U^n(\tau+h) - V_{F(U_\tau^n)}(h)\big)^1 &= \big(\pi_2\big(F\big(U_{\tau_i}^n\big)\big)^1 - \pi_2\big(F\big(U_\tau^n\big)\big)^1\big),\\
h^{-1}\big(U^n(\tau+h) - V_{F(U_\tau^n)}(h)\big)^2 &= \big(F_{Z,\varphi}\big(U_{\tau_i}^n\big) - F_{Z,\varphi}\big(U_\tau^n\big)\big) + 2(\tau-\tau_i)\pi_2\big(Z\big(U_{\tau_i}^n\big)\big)^2\\
&\quad + h\big[\pi_2\big(Z\big(U_{\tau_i}^n\big)\big)^2 - \pi_2\big(Z\big(U_\tau^n\big)\big)^2\big].
\end{aligned}
$$

Letting

$$M^* = \sup_{X\in B_r(X_0)}\left[\tilde{d}_q\big(F(X),0\big), \sup_{\mathcal{D}}\left\{\sum_l \left|\pi_2\big(Z(X)\big)^2_{t_l,t_{l+1}}\right|^{\frac{q}{2}}\right\}^{\frac{2}{q}}\right]$$

and, by using the property that F is also Lipschitz in the q variation sense,

$$\begin{aligned} h^{-1}d_q\big(U^n(\tau+h), V_{F(U^n_\tau)}(h)\big) &\leqslant C_2 d_p(U_{\tau_i}, U_\tau) + 2M^*|\tau-\tau_i| + 2hM^* \\ &\leqslant (C_2+2)M^*|\tau-\tau_i| + 2hM^* \\ &\leqslant (C_2+2)M^*\frac{\varepsilon_n}{(C_1+2M)} + 2hM^*. \end{aligned}$$

Hence,

$$\lim_{h\to 0} h^{-1}d_p\big(U^n(\tau+h), V_{F(U^n_\tau)}(h)\big) = \frac{(C_2+2)M^*}{C_1+2M}\varepsilon_n$$

and so letting $n\to\infty$ gives the result. □

Lemma 28. *If a solution exists, then it is unique.*

Proof. Suppose there exist two solutions U and $\tilde{U}$ of the flow equation with the same initial data. Consider the function

$$g(\tau) := d_q\big(U(\tau), \tilde{U}(\tau)\big).$$

Now,

$$\begin{aligned} \frac{g(\tau+\varepsilon)-g(\tau)}{\varepsilon} \leqslant{}& \frac{d_q(U(\tau+\varepsilon), V_{U(\tau)}(\varepsilon))}{\varepsilon} + \frac{d_q(V_{\tilde{U}(\tau)}(\varepsilon), \tilde{U}(\tau+\varepsilon))}{\varepsilon} \\ &+ \frac{d_q(V_{U(\tau)}(\varepsilon), V_{\tilde{U}(\tau)}(\varepsilon)) - d_q(U(\tau), \tilde{U}(\tau))}{\varepsilon}. \end{aligned}$$

The first two terms on the right-hand side tend to zero as $\varepsilon\to 0$ as U and $\tilde{U}$ are solutions, so let us examine the final term. Recall that

$$\begin{aligned} V_{U(\tau)}(\varepsilon)^1 &= U^1(\tau) + \varepsilon\pi_2\big(F\big(U(\tau)\big)\big)^1, \\ V_{U(\tau)}(\varepsilon)^2 &= U^2(\tau) + \varepsilon F_{Z,\varphi}\big(U(\tau)\big) + \varepsilon^2\pi_2\big(Z\big(U(\tau)\big)\big)^2 \end{aligned}$$

so that by the Lipschitz property of F and the homogeneity of the distance,

$$\begin{aligned} d_q\big(V_{U(\tau)}(\varepsilon), V_{\tilde{U}(\tau)}(\varepsilon)\big) \leqslant{}& d_q\big(U(\tau), \tilde{U}(\tau)\big) + C\varepsilon d_q\big(U(\tau), \tilde{U}(\tau)\big) \\ &+ \varepsilon^2 d_q\big(Z\big(U(\tau)\big), Z\big(\tilde{U}(\tau)\big)\big). \end{aligned}$$

Therefore, $g'(\tau)\leqslant Cg(\tau)$ and an application of Gronwall's inequality gives $g(\tau)=0$.

Combining the above results, we have the following. □

Theorem 29. *If F is a locally Lipschitz near X_0 vector field on $WG\Omega_p$, then there exists a unique solution $U:[0,\alpha]\to\Omega_q(V)$ to the flow equation for $q>p$.*

Remark 30. Using the preceding arguments we would not necessarily have a global solution even for a globally Lipschitz vector field. This is because in our definition we do not assume a Lipschitz type condition on the projection $\pi_2(Z(\cdot))^2$ which appears in the definition of the length of the interval. If however we impose the slightly stronger condition that both the maps Z and φ are Lipschitz, then the induced vector field will be Lipschitz and, moreover, the length of the interval where the solution is defined is larger than a constant which depends only on the Lipschitz constants of Z and φ. In this setting, we have global solutions.

Theorem 31. *Let $Z : WG\Omega_p(V) \to WG\Omega_p(V \oplus V)$ and $\varphi : WG\Omega_p(V) \to WG\Omega_{\frac{p}{2}}(V)$ be maps such that $\pi_1(Z(X)) = X$. If there exist constants C_i for $i = \{1, \dots, 4\}$ such that*

$$d_p\big(Z(X), Z(Y)\big) \leqslant C_1 d_p(X, Y),$$
$$d_{\frac{p}{2}}\big(\varphi(X), \varphi(Y)\big) \leqslant C_2 d_p(X, Y)$$

and

$$d_q\big(Z(X), Z(Y)\big) \leqslant C_3 d_q(X, Y),$$
$$d_{\frac{q}{2}}\big(\varphi(X), \varphi(Y)\big) \leqslant C_4 d_q(X, Y)$$

then there exists a global solution $U : [0, \infty] \to \Omega_q(V)$ to the flow equation with initial data X_0 for the vector field induced by Z and φ for $q > p$.

Proof. If F denotes the vector field induced by Z and φ, the above Lipschitz conditions imply that $\tilde{d}_p(F(X), F(Y)) \leqslant \max[1, C_1 + C_2] d_p(X, Y)$ and $\tilde{d}_q(F(X), F(Y)) \leqslant \max[1, C_3 + C_4] d_q(X, Y)$. Letting

$$C_5 = \max[1, C_1 + C_2],$$
$$C_6 = \max\big[\tilde{d}_p\big(F(X_0), 0\big), d_p\big(Z(X_0), 0\big)\big],$$

we establish the bounds

$$\tilde{d}_p\big(F(X), 0\big) \leqslant C_5 d_p(X, X_0) + \tilde{d}_p\big(F(X_0), 0\big)$$

and

$$\begin{aligned} \sup_{\mathcal{D}} \bigg\{ \sum_l \big|\pi_2\big(Z(X)\big)^2_{t_l, t_{l+1}}\big|^{\frac{p}{2}} \bigg\}^{\frac{2}{p}} &\leqslant d_p\big(Z(X), 0\big) \\ &\leqslant C_1 d_p(X, X_0) + d_p\big(Z(X_0), 0\big). \end{aligned}$$

For $r_1 > 0$, set

$$\begin{aligned} M_1 = & \sup_{X \in B_{r_1}(X_0)} \bigg[\tilde{d}_p\big(F(X), 0\big), \sup_{\mathcal{D}} \bigg\{ \sum_l \big|\pi_2\big(Z(X)\big)^2_{t_l, t_{l+1}}\big|^{\frac{p}{2}} \bigg\}^{\frac{2}{p}} \bigg] \\ & \leqslant C_5 r_1 + C_6 \end{aligned}$$

and apply Lemma 24 so that we have a local ε-solution $U_1^\varepsilon(\cdot)$ on $[0, \alpha_1]$ for $\alpha_1 = \frac{r_1}{2M_1}$. We now apply the lemma again for initial point $U_1^\varepsilon(\alpha_1)$, $r_2 > 0$. To bound M_2, we derive the estimates

$$\begin{aligned} \tilde{d}_p\big(F(X), 0\big) &\leqslant C_5 d_p\big(X, U_1^\varepsilon(\alpha_1)\big) + \tilde{d}_p\big(F\big(U_1^\varepsilon(\alpha_1)\big), 0\big) \\ &\leqslant C_5\big[d_p\big(X, U_1^\varepsilon(\alpha_1)\big) + d_p\big(U_1^\varepsilon(\alpha_1), X_0\big)\big] + \tilde{d}_p\big(F(X_0), 0\big) \end{aligned}$$

and

$$\begin{aligned}\sup_{\mathcal{D}}\Big\{\sum_l \big|\pi_2\big(Z(X)\big)^2_{t_l,t_{l+1}}\big|^{\frac{p}{2}}\Big\}^{\frac{2}{p}} &\leqslant d_p\big(Z(X),0\big)\\ &\leqslant C_1\big[d_p\big(X,U_1^\varepsilon(\alpha_1)\big)+d_p\big(U_1^\varepsilon(\alpha_1),X_0\big)\big]+d_p\big(Z(X_0),0\big).\end{aligned}$$

Therefore, as $d_p(U_1^\varepsilon(\alpha_1),X_0)\leqslant r_1$,

$$\begin{aligned}M_2 &= \sup_{X\in B_{r_2}(U_1(\alpha_1))}\Big[\tilde{d}_p\big(F(X),0\big),\sup_{\mathcal{D}}\Big\{\sum_l\big|\pi_2\big(Z(X)\big)^2_{t_l,t_{l+1}}\big|^{\frac{p}{2}}\Big\}^{\frac{2}{p}}\Big]\\ &\leqslant C_5(r_1+r_2)+C_6\end{aligned}$$

and we get an ε-solution, U_2^ε defined on $[0,\alpha_2]$ where $U_2^\varepsilon(0)=U_1^\varepsilon(\alpha_1)$ and

$$\begin{aligned}\alpha_2 &= \frac{r_2}{2M_2}\\ &\geqslant \frac{1}{2C_5(\frac{r_1+r_2}{r_2})+\frac{C_6}{r_3}}.\end{aligned}$$

If we repeat this process n times, then for each n we obtain an ε-solution, U_n^ε defined on interval of length

$$\alpha_n \geqslant \frac{1}{2C_5(\frac{\sum_{i=1}^{n-1}r_i}{r_n})+\frac{C_6}{r_n}}.$$

We need a lower bound for α_n independent of n. If we choose $r_i=e^i$ then

$$\begin{aligned}\frac{\sum_{i=1}^{n-1}r_i}{r_n} &= e^{1-n}+e^{2-n}+\cdots+e^{-1}\\ &= \sum_{j=1}^{n-1}e^{-j}\end{aligned}$$

and by the ratio test, the series $\sum_{j=1}^{\infty}e^{-j}$ converges to some value C. Then with the above choice of r_i,

$$\begin{aligned}\alpha_n &\geqslant \frac{1}{2C_5(\frac{\sum_{i=1}^{n-1}r_i}{r_n})+\frac{C_6}{r_n}}\\ &\geqslant \frac{1}{2C_5C+\frac{C_6}{e}}.\end{aligned}$$

This means that we can repeat the procedure indefinitely to construct ε-solutions U_n^ε each defined on the interval $[0,\alpha_n]$, with $U_n^\varepsilon(0)=U_{n-1}^\varepsilon(\alpha_{n-1})$ and the total length of the intervals is infinite.

From here, we follow the same arguments as in the local solution case. Take $(\varepsilon_m)_{m\geqslant 0}$ such that $\varepsilon_m\to 0$ as $m\to\infty$. By Lemma 26, each $U_n^{\varepsilon_m}$ has a subsequence which converges in $C([0,\alpha_n],WG\Omega_q(V))$ equipped with the uniform norm to U_n. For U_1^ε, take the convergent subsequence $\{U_1^{\varepsilon_{l_1}}\}$ and for the next step, take a convergent subsequence $\{U_2^{\varepsilon_{l_2}}\}$ of $\{U_2^{\varepsilon_{l_1}}\}$, etc. The indices of the subsequences are given by $\{\varepsilon_{l_1}\} \supseteq \{\varepsilon_{l_2}\} \supseteq \cdots$ so if we take the diagonal subsequence whose indices ε_l are given by $\varepsilon_l = \varepsilon_{l_l}$ we have simultaneous convergence. Since we have uniform, and therefore pointwise convergence, $U_n^{\varepsilon_l}(\tau) \to U_n(\tau)$ for each τ and in particular

$$\begin{aligned} U_n^{\varepsilon_l}(0) &= \lim_{l\to\infty} U_n^{\varepsilon_l}(0) \\ &= \lim_{l\to\infty} U_{n-1}^{\varepsilon_l}(\alpha_{n-1}) \\ &= U_{n-1}(\alpha_{n-1}). \end{aligned}$$

We can put these solutions together to form $\hat{U} : [0,\infty) \to \Omega_q(V)$ defined by

$$\hat{U}(\tau) = \begin{cases} U_1(\tau) & \text{for } \tau \in [0,\alpha_1], \\ U_2(\tau - \alpha_1) & \text{for } \tau \in (\alpha_1, \alpha_1 + \alpha_2], \\ \vdots & \vdots \end{cases}$$

which is then a unique global solution by Lemmas 27 and 28. □

Appendix A

As we make heavy use of the Lyons–Victoir extension, we present here a small extension of the result for paths in $\mathbb{R}^d$. We remark that it may be used to define a mapping from a curve of rough paths $X(\varepsilon)$ to another rough path which can be viewed as $\int X(\varepsilon)\,d\varepsilon$. The construction follows closely the proof of the $\mathbb{R}^d$ case extension in the Lyons–Victoir paper [17].

Let $X(\varepsilon)$ be a path taking values in $WG\Omega_p(\mathbb{R}^d)$. If $X(\varepsilon)$ takes its values only in smooth rough paths then we could define the integral X^μ via the Bochner integral:

$$X_{s,t}^{\mu,1} = \int_0^1 \left[\int_{s<u<t} dx_u(\varepsilon) \right] \mu(d\varepsilon),$$

$$X_{s,t}^{\mu,2} = \int_0^1 \int_0^1 \left[\int_{s<u_1<u_2<t} dx_{u_1}(\varepsilon) \otimes dx_{u_2}(\delta) \right] \mu(d\varepsilon)\,\mu(d\delta).$$

The term in brackets in the definition of $X_{s,t}^{\mu,2}$ can be realized for non-smooth rough paths as $\pi_{1,2}(Z)$ where Z is a rough path in $WG\Omega_p(\mathbb{R}^d \oplus \mathbb{R}^d)$ which extends $(X(\varepsilon), X(\delta))$. Therefore, in order to define the integral, we provide some conditions which allow extension of $(X(\varepsilon), X(\delta))$ to a Z which we can integrate in $T^2(\mathbb{R}^d)$.

Suppose now that we have a family of weakly geometric rough paths $X(\varepsilon)$ with $\varepsilon \in [0,1]$ with associated paths $x(\varepsilon)$ taking values in $\mathbb{R}^d$. We make the following assumptions:

Condition 32. Let $X(\varepsilon)$ for $\varepsilon \in [0,1]$ be a path in $WG\Omega_p(\mathbb{R}^d)$ such that there exists a non-negative, 0 on the diagonal, super additive function $\omega : \Delta_T \to \mathbb{R}$ satisfying

$$\left|X_{s,t}^i(\varepsilon)\right| \leqslant C\omega(s,t)^{\frac{i}{p}}$$

and

$$\left|X_{s,t}^i(\varepsilon) - X_{s,t}^i(\tilde{\varepsilon})\right| \leqslant C|\varepsilon - \tilde{\varepsilon}|\omega(s,t)^{\frac{i}{p}}.$$

Given these conditions, we reformulate in terms of $\frac{1}{p}$ Hölder paths. If none of the paths are constant over any interval, then we can define the bijection $\tau(t) = \omega(0,t)\frac{T}{\omega(0,T)}$ and reparametrize so that

$$\begin{aligned}\left|X^i_{\tau^{-1}(s),\tau^{-1}(t)}\right| &\leqslant C\omega\big(\tau^{-1}(s),\tau^{-1}(t)\big)^{\frac{i}{p}}\\ &\leqslant C\big[\omega\big(\tau^{-1}(0),\tau^{-1}(t)\big)-\omega\big(\tau^{-1}(0),\tau^{-1}(s)\big)\big]^{\frac{i}{p}}\\ &= C\frac{\omega(0,T)^{\frac{i}{p}}}{T}(t-s)^{\frac{i}{p}}. \end{aligned}\tag{A.1}$$

Similarly

$$\left|X^i_{\tau^{-1}(s),\tau^{-1}(t)}(\varepsilon)-X^i_{\tau^{-1}(s),\tau^{-1}(t)}(\tilde{\varepsilon})\right| \leqslant C|\varepsilon-\tilde{\varepsilon}|\frac{\omega(0,T)^{\frac{i}{p}}}{T}(t-s)^{\frac{i}{p}}. \tag{A.2}$$

As a result, we will assume for the rest of the section that we have reparametrized so that the Hölder estimates in (A.1) and (A.2) hold, i.e. when we write $X_{s,t}(\varepsilon)$ it is understood as the reparametrized $X_{\tau^{-1}(s),\tau^{-1}(\tau)}$.

We construct $Z(\varepsilon,\delta)$ in the following manner. For each ε and δ, we define a choice of area element over dyadic intervals associated to the path $(x(\varepsilon),x(\delta))$ in $\mathbb{R}^{2d}$. We then show that over these dyadic points the area element satisfies an estimate of the same type as (A.2). Finally we show that the area element can be extended to all intervals such that the estimate still holds. The constructed $Z(\varepsilon,\delta)$ will then be continuous as a map from $[0,1]^2 \to WG\Omega_p(\mathbb{R}^{2d})$ so we can define the integral of the constituent projection $\pi_{1,2}(Z(\varepsilon,\delta))$.

From the definition of the group imbedding of weakly geometric rough paths into $T^2(\mathbb{R}^{2d})$ we get in the original formulation that $Z(\varepsilon,\delta)^2$ should be defined by

$$Z(\varepsilon,\delta)^2 = \begin{pmatrix} \frac{1}{2}x^i(\varepsilon)x^j(\varepsilon)+A^{ij}(\varepsilon) & \frac{1}{2}x^i(\varepsilon)x^j(\delta)+A^{ij}(\varepsilon,\delta)\\ \frac{1}{2}x^i(\delta)x^j(\varepsilon)+A^{ij}(\varepsilon,\delta) & \frac{1}{2}x^i(\delta)x^j(\delta)+A^{ij}(\delta)\end{pmatrix}$$

where in the $1,1$ entry of the block matrix, $i,j\in\{1,\dots,d\}$, in the $1,2$ entry $i\in\{1,\dots,d\}$ $j\in\{d+1,\dots,2d\}$, etc. Where $A^{ij}_{s,t}$ is interpreted as the area in the (i,j) plane bounded by the curve and its chord between s and t. The case of A^{ij} corresponding to the $1,2$ entry in the block matrix for Z is the only one treated as precisely the same arguments are used in the other terms. In order to be multiplicative, the area elements must satisfy

$$A^{ij}_{s,u} = A^{ij}_{s,t}+A^{ij}_{tu}+\frac{1}{2}\big(x^i_{s,t}(\varepsilon)x^j_{t,u}(\delta)-x^j_{s,t}(\delta)x^i_{t,u}(\varepsilon)\big). \tag{A.3}$$

Lemma 33. *If $X(\varepsilon)$ satisfies Condition* 32 *and is also assumed to be reparametrized then there exists a map A taking the dyadic points of the simplex Δ_T to the set of antisymmetric* 2 *tensors over $\mathbb{R}^{2d}$ satisfying* (A.3) *such that*

$$\begin{aligned}\left|A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)-A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\tilde{\varepsilon},\tilde{\delta})\right| &\leqslant \frac{1}{2}C^2\frac{\omega(0,T)^{\frac{2}{p}}}{T^2}\left[\sum_{l=0}^{n-1}2^{\frac{l(2-p)}{p}}\right]\\ &\quad\times\big(|\varepsilon-\tilde{\varepsilon}|+|\delta-\tilde{\delta}|\big)2^{-\frac{2n}{p}}.\end{aligned}$$

Proof. For each (ε,δ) we can define $A^{ij}(\varepsilon,\delta)$ as follows. Set $A^{ij}_{0,1}(\varepsilon,\delta)=C$ then suppose we have defined $A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)$. We then define A^{ij} on the next finest partition by setting the areas over each of the two halves of the previous partition to be equal. In other words, we set A^{ij} from an old partition point to the point added by the finer partition equal to A^{ij} over the added point to the next old point, i.e. $A^{ij}_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\varepsilon,\delta)=A^{ij}_{\frac{2k+1}{2^{n+1}},\frac{2k+2}{2^{n+1}}}(\varepsilon,\delta)$. Next, as we want the final product to satisfy Eq. (A.3) we set

$$\begin{aligned}A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)&=2A^{ij}_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\varepsilon,\delta)\\&\quad+\frac{1}{2}\bigl(x^i_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\varepsilon)x^j_{\frac{2k+1}{2^{n+1}},\frac{k+1}{2^n}}(\delta)-x^j_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\delta)x^i_{\frac{2k+1}{2^{n+1}},\frac{k+1}{2^n}}(\varepsilon)\bigr)\end{aligned}$$

so that

$$\begin{aligned}A^{ij}_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\varepsilon,\delta)&=A^{ij}_{\frac{2k+1}{2^{n+1}},\frac{2k+2}{2^{n+1}}}(\varepsilon,\delta)\\&=\frac{1}{2}A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)\\&\quad-\frac{1}{4}\bigl(x^i_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\varepsilon)x^j_{\frac{2k+1}{2^{n+1}},\frac{k+1}{2^n}}(\delta)-x^j_{\frac{2k}{2^{n+1}},\frac{2k+1}{2^{n+1}}}(\delta)x^i_{\frac{2k+1}{2^{n+1}},\frac{k+1}{2^n}}(\varepsilon)\bigr).\end{aligned}$$

Then, through the use of induction, we explicitly see what the area is over each of the dyadic points for any level. Indeed

$$A^{i,j}_{0,\frac{1}{2}}(\varepsilon,\delta)=\frac{1}{2}C-\frac{1}{4}\bigl(x^i_{0,\frac{1}{2}}(\varepsilon)x^j_{\frac{1}{2},1}(\delta)-x^j_{0,\frac{1}{2}}(\delta)x^i_{\frac{1}{2},1}(\varepsilon)\bigr)$$

and

$$\begin{aligned}A^{i,j}_{0,\frac{1}{4}}(\varepsilon,\delta)=\frac{1}{4}C-\frac{1}{4}\biggl[&\frac{1}{2}\bigl(x^i_{0,\frac{1}{2}}(\varepsilon)x^j_{\frac{1}{2},1}(\delta)-x^j_{0,\frac{1}{2}}(\delta)x^i_{\frac{1}{2},1}(\varepsilon)\bigr)\\&+\bigl(x^i_{0,\frac{1}{4}}(\varepsilon)x^j_{\frac{1}{4},\frac{1}{2}}(\delta)-x^j_{0,\frac{1}{4}}(\delta)x^i_{\frac{1}{4},\frac{1}{2}}(\varepsilon)\bigr)\biggr]\end{aligned}$$

so that

$$\begin{aligned}&A^{i,j}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)\\&\quad=\frac{1}{2^n}C-\frac{1}{4}\sum_{l=0}^{n-1}2^{-l}\bigl(x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\varepsilon)x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\delta)-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\delta)x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\varepsilon)\bigr),\end{aligned}$$

for all $n\geqslant 1$. Therefore,

$$\begin{aligned}&A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)-A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\tilde{\varepsilon},\tilde{\delta})\\&\quad=-\frac{1}{4}\sum_{l=0}^{n-1}2^{-l}\Bigl(\bigl(x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\varepsilon)x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\delta)-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\delta)x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\varepsilon)\bigr)\\&\qquad-\bigl(x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\tilde{\varepsilon})x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde{\delta})-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\tilde{\delta})x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde{\varepsilon})\bigr)\Bigr)\end{aligned}$$

and

$$\begin{aligned}&\big|A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)-A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\tilde\varepsilon,\tilde\delta)\big|\\&\leqslant\frac14\left|\sum_{l=0}^{n-1}2^{-l}\Big[\Big(x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\varepsilon)x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\delta)-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\delta)x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\varepsilon)\Big)\right.\\&\qquad\left.-\Big(x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\tilde\varepsilon)x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\delta)-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\tilde\delta)x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\varepsilon)\Big)\Big]\right|.\end{aligned}$$

From here we can add and subtract $x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\varepsilon)x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\delta)$ and $x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\delta)x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\varepsilon)$ to establish that the term in square brackets above is equal to

$$\begin{aligned}&\Big[\Big(x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\varepsilon)\Big[x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\delta)-x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\delta)\Big]\\&\quad-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\delta)\Big[x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\varepsilon)-x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\varepsilon)\Big]\Big)\\&\quad-\Big(\Big[x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\tilde\varepsilon)-x^i_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\varepsilon)\Big]x^j_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\delta)\\&\quad\times\Big[x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\delta)-x^j_{\frac{2k}{2^{n-l}},\frac{2k+1}{2^{n-l}}}(\tilde\delta)\Big]x^i_{\frac{2k+1}{2^{n-l}},\frac{2k+2}{2^{n-l}}}(\tilde\varepsilon)\Big)\Big].\end{aligned}$$

Hence by using our conditions, we have

$$\begin{aligned}\big|A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)-A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\tilde\varepsilon,\tilde\delta)\big|&\leqslant\frac12C^2\frac{\omega(0,T)^{\frac2p}}{T^2}\left[\sum_{l=0}^{n-1}2^{\frac{l(2-p)}{p}}\right]\\&\quad\times\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)2^{-\frac{2n}{p}}.\qquad\square\end{aligned}$$

Now, since we currently have only the extension defined for dyadic points of time, let us show that we can extend the area element to all points of the simplex Δ_T such that the Hölder estimate still holds.

Lemma 34. *There exists a unique $\hat{A}^{ij}$ defined on all points of Δ_T which on dyadic points coincides with A^{ij} defined above such that*

$$\big|\hat{A}^{ij}_{s,t}(\varepsilon,\delta)-\hat{A}^{ij}_{s,t}(\tilde\varepsilon,\tilde\delta)\big|\leqslant\tilde{C}\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)(t-s)^{\frac2p}$$

for all $s,t\in[0,1]$ for a constant $\tilde{C}$ depending on p.

Proof. This proof follows the proof of Lemma 2 in the Lyons–Victoir extension paper [17]. We have established that, after reparametrization, the area elements given above on dyadic points satisfy

$$\big|A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\varepsilon,\delta)-A^{ij}_{\frac{k}{2^n},\frac{k+1}{2^n}}(\tilde\varepsilon,\tilde\delta)\big|\leqslant4\hat{C}\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)2^{-\frac{2n}{p}}$$

where $\hat{C}=\frac12C^2\frac{\omega(0,T)^{\frac2p}}{T^2}[\sum_{l=0}^{\infty}2^{\frac{l(2-p)}{p}}]$, since $\sum_{l=0}^{n-1}2^{\frac{l(2-p)}{2}}$ converges as $2<p<3$. The first step is to prove the second inequality when s,t are dyadic points from the same level of fineness but not necessarily consecutive points. Let $D_m=\bigcup_{k=0}^{2^m}\frac{k}{2^m}$ and consider all $s,t\in D_m$ such that

$0 < t - s < 2^{-r}$ for some fixed integer r. We want to show by induction that

$$\left|A_{s,t}^{ij}(\varepsilon,\delta) - A_{s,t}^{ij}(\tilde{\varepsilon},\tilde{\delta})\right| \leqslant 4\hat{C}\left(|\varepsilon-\tilde{\varepsilon}| + |\delta-\tilde{\delta}|\right)\sum_{k=r+1}^{m} 2^{-\frac{2k}{p}}.$$

For the case when $m = r+1$ the above is identical to the condition. For the inductive step, assume the statement is true for $m = r+1,\dots,M-1$ and consider $s,t \in D_M$ with $0 < t - s < 2^{-r}$. Define two points s_1 and t_1 from the next coarsest level of dyadic points which are nonetheless adjacent to s and t, i.e. $s_1 = \min\{u \in D_{M-1}\colon u \geqslant s\}$ and $t_1 = \max\{u \in D_{M-1}\colon u \leqslant t\}$. Then

$$\begin{aligned}
&\left|A_{s,s_1}^{ij}(\varepsilon,\delta) - A_{s,s_1}^{ij}(\tilde{\varepsilon},\tilde{\delta})\right| \leqslant \hat{C}\left(|\varepsilon-\tilde{\varepsilon}| + |\delta-\tilde{\delta}|\right)2^{-\frac{2M}{p}},\\
&\left|A_{t_1,t}^{ij}(\varepsilon,\delta) - A_{t_1,t}^{ij}(\tilde{\varepsilon},\tilde{\delta})\right| \leqslant \hat{C}\left(|\varepsilon-\tilde{\varepsilon}| + |\delta-\tilde{\delta}|\right)2^{-\frac{2M}{p}}.
\end{aligned}$$

Now, the area elements A^{ij} corresponding to the $1,2$ entry in the block matrix for Z were constructed to satisfy

$$A_{s,t}^{ij}(\varepsilon,\delta) = A_{s,s_1}^{ij}(\varepsilon,\delta) + A_{s_1,t}^{ij}(\varepsilon,\delta) + \frac{1}{2}\left(x_{s,s_1}^i(\varepsilon)x_{s_1,t}^j(\delta) - x_{s,s_1}^j(\varepsilon)x_{s,t}^i(\delta)\right)$$

so that

$$\begin{aligned}
&A_{s,t}^{ij}(\varepsilon,\delta) - A_{s,t}^{ij}(\tilde{\varepsilon},\tilde{\delta})\\
&\quad= A_{s,s_1}^{ij}(\varepsilon,\delta) - A_{s,s_1}^{ij}(\tilde{\varepsilon},\tilde{\delta}) + A_{s_1,t_1}^{ij}(\varepsilon,\delta) - A_{s_1,t_1}^{ij}(\tilde{\varepsilon},\tilde{\delta}) + A_{t_1,t}^{ij}(\varepsilon,\delta) - A_{t_1,t}^{ij}(\tilde{\varepsilon},\tilde{\delta})\\
&\qquad+ \frac{1}{2}\left(x_{s,s_1}^i(\varepsilon)x_{s_1,t}^j(\delta) - x_{s,s_1}^j(\varepsilon)x_{s,t}^i(\delta)\right) - \frac{1}{2}\left(x_{s,s_1}^i(\tilde{\varepsilon})x_{s_1,t}^j(\tilde{\delta}) - x_{s,s_1}^j(\tilde{\varepsilon})x_{s,t}^i(\tilde{\delta})\right)\\
&\qquad+ \frac{1}{2}\left(x_{s_1,t_1}^i(\varepsilon)x_{t_1,t}^j(\delta) - x_{s_1,t_1}^j(\delta)x_{t_1,t}^i(\varepsilon)\right) - \frac{1}{2}\left(x_{s_1,t_1}^i(\tilde{\varepsilon})x_{t_1,t}^j(\tilde{\delta}) - x_{s_1,t_1}^j(\tilde{\delta})x_{t_1,t}^i(\tilde{\varepsilon})\right).
\end{aligned}$$

Applying the inductive step and adding and subtracting $x_{s,s_1}^i(\varepsilon)x_{s_1,t}^j(\tilde{\delta})$, $x_{s,s_1}^j(\delta)x_{s_1,t}^i(\tilde{\varepsilon})$, $x_{s_1,t_1}^i(\varepsilon)x_{t_1,t}^j(\tilde{\delta})$, and $x_{s_1,t_1}^j(\delta)x_{t_1,t}^i(\tilde{\varepsilon})$ we get

$$\begin{aligned}
&\left|A_{s,t}^{ij}(\varepsilon,\delta) - A_{s,t}^{ij}(\tilde{\varepsilon},\tilde{\delta})\right|\\
&\quad\leqslant 2\hat{C}\left(|\varepsilon-\tilde{\varepsilon}| + |\delta-\tilde{\delta}|\right)2^{-\frac{2M}{p}} + \hat{C}\left(|\varepsilon-\tilde{\varepsilon}| + |\delta-\tilde{\delta}|\right)\sum_{k=r+1}^{M-1} 2^{-\frac{2k}{p}}\\
&\qquad+ \frac{1}{2}\left|x_{s,s_1}^i(\varepsilon)\left[x_{s_1,t}^j(\delta) - x_{s_1,t}^j(\tilde{\delta})\right]\right| + \frac{1}{2}\left|x_{s,s_1}^j(\delta)\left[x_{s_1,t}^i(\varepsilon) - x_{s_1,t}^i(\tilde{\varepsilon})\right]\right|\\
&\qquad+ \frac{1}{2}\left|\left[x_{s,s_1}^i(\tilde{\varepsilon}) - x_{s,s_1}^i(\varepsilon)\right]x_{s_1,t}^j(\tilde{\delta})\right| + \frac{1}{2}\left|\left[x_{s,s_1}^j(\tilde{\delta}) - x_{s,s_1}^j(\delta)\right]x_{s,t}^i(\tilde{\varepsilon})\right|\\
&\qquad+ \frac{1}{2}\left|x_{s_1,t_1}^i(\varepsilon)\left[x_{t_1,t}^j(\delta) - x_{t_1,t}^j(\tilde{\delta})\right]\right| + \frac{1}{2}\left|x_{s_1,t_1}^j(\delta)\left[x_{t_1,t}^i(\varepsilon) - x_{t_1,t}^i(\tilde{\varepsilon})\right]\right|\\
&\qquad+ \frac{1}{2}\left|\left[x_{s_1,t_1}^i(\tilde{\varepsilon}) - x_{s_1,t_1}^i(\varepsilon)\right]x_{t_1,t}^j(\tilde{\delta})\right| + \frac{1}{2}\left|\left[x_{s_1,t_1}^j(\tilde{\delta}) - x_{s_1,t_1}^j(\delta)\right]x_{t_1,t}^i(\tilde{\varepsilon})\right|.
\end{aligned}$$

Finally, we apply the condition to get

$$\begin{aligned}\left|A^{ij}_{s,t}(\varepsilon,\delta)-A^{ij}_{s,t}(\tilde\varepsilon,\tilde\delta)\right| &\leqslant 2\hat C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)2^{-\frac{2M}{p}}\\ &\quad+4\hat C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)\sum_{k=r+1}^{M-1}2^{-\frac{2k}{p}}\\ &\quad+2\hat C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)2^{-\frac{2M}{p}}\end{aligned}$$

so that

$$\left|A^{ij}_{s,t}(\varepsilon,\delta)-A^{ij}_{s,t}(\tilde\varepsilon,\tilde\delta)\right| \leqslant 4\hat C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)\sum_{k=r+1}^{M}2^{-\frac{2k}{p}}$$

as required. This works because the sum $\sum_{l=0}^{\infty}2^{\frac{l(2-p)}{p}}$ is larger than 1. For all points $s,t\in\bigcup_m D_m$ such that $2^{-(r+1)}<t-s<2^{-r}$

$$\begin{aligned}\left|A^{ij}_{s,t}(\varepsilon,\delta)-A^{ij}_{s,t}(\tilde\varepsilon,\tilde\delta)\right| &\leqslant 4\hat C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)\sum_{k=r+1}^{\infty}2^{-\frac{2k}{p}}\\ &=4\hat C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)2^{\frac{-2(r+1)}{p}}\sum_{k=0}^{\infty}2^{-\frac{2k}{p}}\\ &\leqslant \tilde C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)(t-s)^{-\frac{2}{p}}.\end{aligned}$$

Next, using the fact that for any real number t, $\frac{\lfloor 2^r t\rfloor}{2^r}\to t$ as $r\to\infty$, we define for arbitrary $s,t\in[0,T]$

$$\hat A^{ij}_{s,t}(\varepsilon,\delta)=\lim_{r\to\infty}A^{ij}_{\frac{\lfloor 2^r s\rfloor}{2^r},\frac{\lfloor 2^r t\rfloor}{2^r}}(\varepsilon,\delta)$$

and so, by continuity of the norm, the required estimate holds. $\square$

From this we can define an extension $Z(\varepsilon,\delta)$ of $(X(\varepsilon),X(\delta))$ where

$$Z(\varepsilon,\delta)^2=\begin{pmatrix}\frac12 x^i(\varepsilon)x^j(\varepsilon)+\hat A^{ij}(\varepsilon) & \frac12 x^i(\varepsilon)x^j(\delta)+\hat A^{ij}(\varepsilon,\delta)\\ \frac12 x^i(\delta)x^j(\varepsilon)+\hat A^{ij}(\varepsilon,\delta) & \frac12 x^i(\delta)x^j(\delta)+\hat A^{ij}(\delta)\end{pmatrix}.$$

This extension then satisfies

$$\left|Z^2(\varepsilon,\delta)_{s,t}-Z^2(\tilde\varepsilon,\tilde\delta)_{s,t}\right|\leqslant C\big(|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|\big)(t-s)^{-\frac{2}{p}}$$

which means we can bound $d_p(Z(\varepsilon,\delta),Z(\tilde\varepsilon,\tilde\delta))$ by a constant multiplied by $|\varepsilon-\tilde\varepsilon|+|\delta-\tilde\delta|$.

Since we now know the constructed $Z(\varepsilon,\delta)$ is continuous as a map from $(\varepsilon,\delta)\to WG\Omega_p(\mathbb{R}^d)$, it is also measurable and so integration makes sense.

Definition 35. Given a path $X(\varepsilon)$ in $W\Omega_p(\mathbb{R}^d)$ satisfying Condition 32 and a measure μ supported on [0, 1], we define its integral X^μ by

$$X^{\mu,1}_{s,t}=\int_0^1 X^1_{s,t}(\varepsilon)\,\mu(d\varepsilon),$$

$$X^{\mu,2}_{s,t}=\int_0^1\int_0^1 \pi_{1,2}\big(Z(\varepsilon,\delta)\big)_{s,t}\,\mu(d\varepsilon)\,\mu(d\delta)$$

where Z is defined by the extension constructed above.

In the following proposition we show the integrated path is still a rough path.

Proposition 36. *The object X^μ in Definition* 35 *is multiplicative.*

Proof. We have

$$\begin{aligned}
\left(X^\mu_{s,t} \otimes X^\mu_{t,u}\right)^1 &= \int_0^1 X^1_{s,t}(\varepsilon)\,\mu(d\varepsilon) + \int_0^1 X^1_{t,u}(\varepsilon)\,\mu(d\varepsilon) \\
&= \int_0^1 \left[X^1_{s,t}(\varepsilon) + X^1_{t,u}(\varepsilon)\right]\mu(d\varepsilon) \\
&= \int_0^1 X^1_{s,u}(\varepsilon)\,\mu(d\varepsilon) \\
&= X^{\mu,1}_{s,u}
\end{aligned}$$

since X is multiplicative. Also,

$$\begin{aligned}
\left(X^\mu_{s,t} \otimes X^\mu_{t,u}\right)^2 = {}& \int_0^1 \int_0^1 \pi_{1,2}\big(Z(\varepsilon,\delta)\big)_{s,t}\,\mu(d\varepsilon)\,\mu(d\delta) + \int_0^1 \int_0^1 \pi_{1,2}\big(Z(\varepsilon,\delta)\big)_{t,u}\,\mu(d\varepsilon)\,\mu(d\delta) \\
& + \int_0^1 X^1_{s,t}(\varepsilon)\,\mu(d\varepsilon) \otimes \int_0^1 X^1_{t,u}(\delta)\,\mu(d\delta)
\end{aligned}$$

and since $Z(\varepsilon,\delta)$ is multiplicative

$$\pi_{1,2}\big(Z(\varepsilon,\delta)\big)_{s,u} = \pi_{1,2}(Z)_{s,t} + \pi_{1,2}(Z)_{t,u} + X^1_{s,t}(\varepsilon) \otimes X^1_{t,u}(\delta).$$

This together with the linearity of the integrals implies

$$\begin{aligned}
\left(X^\mu_{s,t} \otimes X^\mu_{t,u}\right)^2 &= \int_0^1 \int_0^1 \pi_{1,2}\big(Z(\varepsilon,\delta)\big)_{s,u}\,\mu(d\varepsilon)\,\mu(d\delta) \\
&= X^{\mu,2}_{s,u}. \qquad \square
\end{aligned}$$

References

[1] R.H. Cameron, W.T. Martin, Transformations of Wiener integrals under translations, Ann. of Math. 45 (1944) 386–396.

[2] T. Cass, P. Friz, N. Victoir, Non-degeneracy of Wiener functionals arising from rough differential equations, Trans. Amer. Math. Soc. 361 (2009) 3359–3371.

[3] T. Cass, P. Friz, Densities for rough differential equations under Hörmander's condition, Ann. of Math. 171 (2010) 2115–2141.

[4] F. Cirpriano, A.B. Cruziro, Flows associated to tangent processes on the Wiener space, J. Funct. Anal. 166 (2) (1999) 310–331.

[5] A.B. Cruzeiro, Equations differentielles sur l'espace de Wiener et formules de Cameron–Martin non lineares, J. Funct. Anal. 54 (1983) 206–227.

[6] A.B. Cruziero, P. Malliavin, Renormalized differential geometry on path space: structural equation, curvature, J. Funct. Anal. 139 (1) (1996) 119–181.

[7] B. Driver, A Cameron–Martin type quasi-invariance theorem for Brownian motion on a compact Riemannian manifold, J. Funct. Anal. 110 (1992) 273–376.

[8] B. Driver, Towards calculus and geometry on path spaces, Proc. Sympos. Pure Math. 57 (1995) 405–422.

[9] D. Elworthy, X.M. Li, A class of integration by parts formula in stochastic analysis I, in: N. Ikeda, S. Watanabe, M. Fukushima, H. Kunia (Eds.), Itô's Calculus and Probability Theory, Springer-Verlag, Tokyo, 1996, pp. 15–30.

[10] D. Elworthy, X.M. Li, Itô maps and analysis on path spaces, Math. Z. 257 (3) (2007) 643–706.

[11] P. Friz, N. Victoir, A note on the notion of geometric rough paths, Probab. Theory Related Fields 136 (2006) 395–416.

[12] Y. Hu, A.S. Üstünel, M. Zakai, Tangent processes on Wiener space, J. Funct. Anal. 192 (1) (2002) 234–270.

[13] T. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoamericana 14 (1998) 215–310.

[14] T. Lyons, Z. Qian, A class of vector fields on path spaces, J. Funct. Anal. 145 (1997) 205–223.

[15] T. Lyons, Z. Qian, Flow equations on spaces of rough paths, J. Funct. Anal. 149 (1997) 135–159.

[16] T. Lyons, Z. Qian, System control and rough paths.

[17] T. Lyons, N. Victoir, An extension theorem to rough paths, Ann. Inst. H. Poincaré Anal. Non Linéaire 24 (5) (2007) 835–847.

[18] P. Malliavin, Stochastic calculus of variation and hypoelliptic operators, in: Proc. Intern. Symp. SDE, Kyoto, 1976, Wiley, New York, 1978, pp. 195–263.

[19] P. Malliavin, Naturality of quasi-invariance of some measures, in: Stochastic Analysis and Applications: Proc. Lisbon, 1989, in: Progr. Probab., vol. 26, Birkhäuser, Boston, 1991, pp. 144–154.

[20] M.-P. Malliavin, P. Malliavin, Integration on loop groups I: quasi invariant measures, quasi invariant integration on loop groups, J. Funct. Anal. 93 (1990) 207–237.